DOCUMENT RESUME

ED 204 159 SE 035 267

AUTHOR	Dienes, Zoltan P.: And Others
TITLE	Mathematics Through Thinking: How and Why. Volume I - Learning Modern Mathematics: Principles and Pedagogy. Volume II - Learning Modern Mathematics: Evaluation.
PUB DATE	80
NOTE	739p.: Not available in hard copy due to marginal legibility of original document. Prepared through the International Study Group for Mathematics Learning.
EDRS PRICE	MF04 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Cognitive Development: Curriculum Development: Curriculum Evaluation: Educational Research: Elementary Education: *Elementary School Mathematics: *Learning Theories: Mathematics Curriculum: *Mathematics Instruction: *Program Descriptions: Program Design: Program Evaluation: *Teaching Methods
IDENTIFIERS	*International Study Group for Math Learning: *Mathematics Education Research

ABSTRACT

Presented are new principles and methods in the instruction of mathematics, with the primary focus at the elementary school level. The first half has the following sections: (1) The Influence of Environment on Child Behavior: (2) Pre-Mathematics: (3) The General and the Abstract: (4) The Passage from Pre-Math to Mathematics: (5) Building a Mathematics Program: (6) New Mathematics Programs: (7) The OPI Mathematics Project in Hungary: (8) Dialogue at Sherbrooke, Quebec: (9) A Visitor Looks at Sherbrooke: and (10) The New York Project. The second half is an evaluation of the three studies that are presented in the first. These investigations are referred to as the Budapest, Sherbrooke, and New York Studies. Chapter titles in this portion of the document are: (11) The Nature of Evaluation: (12) Difficulties in Educational Evaluations: (13) Theoretical Model of the Evaluation for the Dienes Programs: (14) The Structure of the Three Evaluation Studies: (15) Traditional and Modern Mathematics: (16) Fatigue (Budapest Study): (17) Attention (Budapest Study): (18) Learning Ability: (19) Creativity: (20) Perservation and Sociometric Structure: (21) Attitude: (22) Anxiety--Achievement Motivation--Effects on Personality: (23) Attitude Toward Problems: and (24) Summary. (MP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED204159

MATHEMATICS THROUGH THINKING

HOW AND WHY

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R.M.C. LANT

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

International Study Group for Mathematics Learning

1980

SE 035 267

VOLUME I

LEARNING MODERN MATHEMATICS - PRINCIPLES AND PEDAGOGY

by

Zoltan P. Dienes

Sometime Director of the Centre for Research in Psychomathematics

University of Sherbrooke

Quebec, Canada

Sándor Klein

Assistant Professor of Psychology

Teacher Training and Pedagogical Institute

Technical University of Budapest, Hungary

Rainer Lant

Associate Professor of Mathematics

Memorial University

St. John's, Newfoundland, Canada

CONTENTS

VOLUME I

Chapter		Page
1.	The Influence of Environment on Child Behaviour	1
2.	Pre-Mathematics	24
3.	The General and the Abstract	44
4.	The Passage from Pre-Math to Mathematics	64
5.	Building a Mathematics Program	80
6.	New Mathematics Programs	124
7.	The OPI Mathematics Project in Hungary	154
8.	Dialogue at Sherbrooke, Quebec.	190
9.	A Visitor looks at Sherbrooke	215
10.	The New York Project	253
	Bibliography	271

Chapter 1

The Influence of Environment on Child Behaviour

1.1 The influence of motivation on behaviour

It is well known by those who have studied the problem that learning is an adaptation process. A living organism reacts to the environment so as to adapt to it; he experiences certain needs with regard to this environment; to satisfy these needs he must develop ways of proceeding; some will succeed, others will prove less suited to the objectives of the organism in the adaptation process. When an organism passes from one form of behaviour, which does not answer his needs well, to a more efficient form, it is said that the organism has learned something. Let us quote an example: most dogs run away at the sight of a person picking up a stone because they know from experience that very often this stone will be aimed at them. Consequently for them they had best dodge the blow; running away in this case seems to be the only possible way out. It is a negative adaptation; the organism wants to avoid an unpleasant happening. There are also positive adaptations. Dogs will smell meat cooking in the kitchen; they will head for the kitchen and demonstrate their wish to share the meal. The smell in such a case is the stimulus to which the dogs respond.

Avoiding the unpleasant and seeking what is pleasant provides motivation for behaviour. When an organism avoids what is unpleasant to satisfy his needs it is said that this organism behaves in a manner adapted to his environment. It is by such an adaptation that a child learns to behave in a manner acceptable in his social environment. Excesses are not tolerated and proper conduct is rewarded. Thus a child will gradually learn to conform to a social discipline by undergoing a process allowing him to satisfy most of his needs while minimising unpleasant outcomes.

We have seen that there exists positive and negative motivations favouring behaviour. Life compels us to behave according to these motivations. On the one hand, the organism acts or does not act the same way, depending on the specified objective aimed at; on the other hand it can choose an activity which allows it to prevent certain likely unpleasant consequences of certain kinds of activity. Most of us work for a salary or some of us work because we find our occupation interesting and possibly more pleasant than another, even if the latter might be more profitable. In both cases, the objective appears acceptable; but after a week's work, we realize that our paycheck is also a vital necessity; we likewise choose among the consequences and our preferance for one or the other of these consequences influences our behaviour.

1.2 Motivation

Among positive motivations, there are some which are primary and others which are secondary. Primary motivation is a desire to attain a definite aim, where the very action which gives us the possibility to attain it is already contained in the realization of this aim. We ski, we like the fresh mountain air and all that is implied in skiing. No other secondary aim follows this activity or sport. In such cases, it is said that motivation is primary, that is, inherent in the action itself. Motivation will be secondary in the case of any activity leading to a consequence not inherent in the action itself. For example, we accomplish our duty at the office or at the factory in order to receive a salary at the end of the week; therefore we do not work because work contains by its very nature the objective aimed at, but because we hope for the financial remuneration derived from it.

A priori, it is evident that primary motivation is preferable to secondary motivation, because the activity chosen contains in itself the desired effect. It follows that its effects are enjoyed during and after the period of time of the action involved. Primary motivation gives us a motive force which supplies energy during the whole action, whereas our paycheck at the end of the week without taking too much trouble, our work might be neglected; after all, to take our ease might appear to us more pleasant than to work hard. Therefore in the case of secondary motivation we are, up to a certain point, in a contradictory situation;

we accomplish something reluctantly and we do it only to reach our objectives.

It is evident that to ensure a behaviour likely to result in reaching a pedagogical objective, it is better to use primary rather than secondary motivation. Motivation in the child should if possible, form an integral part of the action itself. Such motivation is said to be intrinsic. If, for the child, the activity has an intrinsic interest, we can be sure that even without controlling the activity, it will continue under its own steam; however, if we use secondary motivation, that is, extrinsic ones, the child will have to be closely supervised so that he is sure to continue the activity demanded of him. This is what we find in most cases where classical educational methods of punishment and reward are applied. The child is "negatively" motivated if for example, he is forced to study Latin, under threat of punishment; he is said to be "positively" motivated if he wants the approval of his parents or of his teacher, or to obtain a good grade. The system of grades uses essentially an extrinsic type of motivation. Grades are not relevant to the activity we wish to call into play; learning is nothing inherently to do with grades. If we do not use intrinsic motivation, which continually provides the necessary energy for the desired activity, then much energy will be spent by the teacher to keep the activity going. In general, extrinsic motivation aims at a final objective which is too distant for the child to take seriously enough, especially if at the time he is stimulated by other conflicting forms of motivation!

1.3 The role of curiosity

Curiosity is a motive power which governs a number of human activities. Curiosity has induced us to discover all the parts of our globe, to explore space, to probe the secrets of biology, chemistry and the atom, etc... There are other motive powers, for example of a military, social and financial nature. The fundamental motivation of scientists engaged in research is of a primary kind, awakened by the intrinsic curiosity of man for the complexities of his natural surroundings.

A child is curious by nature. He unceasingly scrutinizes his surroundings and this exploration is what we observe as play. Play for him is a means of adapting himself to his surroundings. This adaptation is done, at least partially, through the satisfaction of the instinct of curiosity present in each of us. Even animals to a certain extent have this "instinct". When you move into a new apartment your cat will spend some time examining the nooks and crannies of a room; the children will explore the same corners, besides exploring the immediate neighbourhood: the garden, the streets, the stores round the corner, the nearest cinema, etc... The organism needs to know its environment because it is the dialectical process between the organism and its environment which largely makes up its life.

If something happens which cannot be explained by means of known patterns, the organism then becomes more curious. For example, a child sees a squirrel in the garden and notices that it makes its appearance at a certain hour of the day. The curiosity of the child is awakened and from then on he will want to know the why and the wherefore of this little animal's behaviour. He can imagine the squirrel lives close by in a small hole or that a "friend" pays it a visit. But, before he is able to form a mental image of what is going on, he will spend some time in observation in order to satisfy his curiosity.

When a system of events is quite complex, it becomes much more difficult to make up a pattern in terms of which to explain it and to foresee its further developments. If the events are too complex for him, the child will lose heart and simply admit that he does not know how things happen. If he is capable of it, he will want to form a mental image in order to foresee how things will or are likely to happen. A little girl will watch her mother prepare the meals in the kitchen. There are certain mixes which she prefers to others, for example, the cake mix: she will pay more attention to the succession of movements leading to the making of a cake, in preference to another succession of movements which does not lead to such interesting results. She will interfere with her mother's work, ("let me put it in the oven") and after much observation, she will be in a position to carry out the same operations by herself. This is how she learns the regularity of events leading to a specific result. To make this cake, she learns when the ingredients must be added and blended with one another, because if the operations did not follow one another in a prescribed order, they would not result in a cake. Following a certain number of successes and failures, the method used in making a cake will have quickly taken shape in her mind. The little girl will cook because she loves to cook, because she likes being with her mother and because she also likes the final result, the tasting of the cake. This is an example of an intrinsic motivation, which encourages the child to continue in its activity.

1.4 The role of choice

The pedagogue is involved in the problem of finding situations, such that if a child is placed within them, the result would be the achievement of the aim desired by the pedagogue for the child. Supposing that we wish the child to learn a certain mathematical topic. One effective way to achieve our aim would be to arrange a series of activities which would motivate him in an intrinsic manner. The child would take pleasure in pursuing them, all the while getting nearer the results hoped for. In short, the problem seems to be to find ways to use intrinsic motivation to the maximum in order to create an effective learning situation leading the child in the desired direction. The problem is very simple to work out: the child could be placed in an environment in which he can undertake certain activities which will be judged interesting by him and which he will therefore perform. We must admit that, pedagogically speaking, this amounts to cheating, because the child's choice will not really be a true and free choice,

even though it will have been suggested by the pedagogue. Nevertheless, the child will feel free as far as he is concerned; he will appreciate the decision he has made in exercising a preference. It is always better for the child to feel that the choice was his. It is a truism to say that we accomplish a task more rapidly and more efficiently if we have chosen it ourselves than if it has been imposed. Evidently if the child makes another choice, and if it is to be a real choice, he has to be left free to make it. It would be a mistake for the teacher to create situations and choices which he does not wish the child to take advantage of. We need to be intellectually honest to ourselves as teachers. All possible permissible choices should have useful and interesting built-in sequels. It is clearly not possible for a child to make his choices in absolute liberty; if that were the case, he would learn much better in the street or in the fields, and the teacher would be out of work. To educate a child implies an intervention in its behaviour on the part of the teacher. The environment must nevertheless be arranged so that there is as much choice as possible and so that the alternatives are organized in such a manner as to really allow the child to make choices. If he chooses a certain alternative there must be an interesting and useful sequel to it; if he chooses another, a suitable sequel should equally be provided for this other alternative.

It must be specified that any particular choice, although planned by the teacher, does not necessarily have to be made at any given precise moment. For example, the child might be given a choice between painting, reading, or doing some work in geometry; if the child wishes to read a story rather than study geometry, he must be left free to do so. However, the problem which has been presented to him must be interesting enough for him to come back after a certain time and show interest, and say to the teacher: "Sir, I would now like to do the geometry problem you gave me." After having worked for half an hour on his geometry problem, he might well be mentally tired; he could then relax by doing a little painting. This does not mean that the geometry problem has been abandoned. It is most important for everyone to take a few minutes of relaxation: taking a coffee-break, for example, going on vacation, or taking a weekend of rest after having worked under tense conditions. To regulate working and relaxation periods is a very individual matter. It would be very difficult to impose on the whole class the same amount of work and for it to be maximally efficient for all the children. For this very reason, an environment should be created in the classroom where, effectively, there would exist a choice of activities: some of these would be intellectually tiring but at the same time satisfying the child's intellectual curiosity; other activities such as writing or painting would always have a possibility to relax built into them. If the environment is not organized in this way, the child will always find ways of certain escape. The need to go to the toilet, the feigned headaches and stomach aches constitute unbeatable escape mechanisms through which to avoid the activity which does not interest him any more. If the children resort to none of these excuses to escape from their work, the teacher can be reasonably sure that he has not made any fundamental educational mistakes. However, if the children must resort to all kinds of excuses to get around the tasks assigned to them by the teacher, he can be sure that the situation in the classroom is far from ideal from the educational standpoint.

1.5 Mathematics laboratory. Non-structured material

All that has been just said amounts to the setting up of a laboratory in the classroom. In the case of mathematics, it means the construction of a mathematics laboratory. A laboratory, by definition, is a place of research where problems are encountered and where the necessary material to solve them is at hand. Consequently, much material is needed, accompanied by many suggested problems or situations implying problems by the nature of the material provided.

What kind of material should be provided in the classroom? Essentially, there are two kinds. Different types of objects to be found in the child's surroundings can be introduced into the classroom; for example: sticks, corks or caps, boxes, autumn leaves, rocks, shells, etc... It is of interest to introduce these objects with which the children come into contact daily: this can convince them that there is a relationship between what they see around them, in their immediate surroundings and what they do in school. Too often, especially in the mathematics class, what goes on is devoid of meaning for the child because he does not find any link between what he does in the mathematics class and what goes on in his personal life. If, from his home and his general environment, objects familiar to him are presented for study and discussion, he will see that what he learns in school is related to everyday life and to the problems which he meets with at home, in his village or in his town.

Such material can be used to construct categories, to learn to recognize specific properties as well as the relevance or irrelevance of certain properties in different situations. For example, let us say that the children have brought autumn leaves to the class; pretty pictures could be made out of them by sticking them on sheets of cardboard. Such activities enrich children's ideas of nature and also give them a chance to experience geometry, because in the practical sense they are faced with the problem of distributing their leaves on a limited surface which has a specific shape. But games of categorization could equally be played.

For example, the teacher might say to a group of students:

"Before sticking anything on paper to make pretty pictures out of them, we should know our leaves and assemble those which are alike. We will make three separate piles out of them: the yellow ones, the green ones and the red ones."

There will evidently be discussions as to whether a leaf is red or yellow. How can one decide if a leaf is red or yellow? The children will understand that being red is not a specific property of a leaf, but if we say " Here is a leaf having some red in it" , we are speaking of a specific property..

In this case, we have agreed on what it is to "have some red in it". When the children have understood that "red" is not sufficient, but that they must say:

"There is red in the leaf" to specify the property of the leaf which allows us to put it in a pile, they will have grappled with a fundamental problem.

Alternatively one could say:

"This leaf is all red" or

"The other leaf is not entirely red, because there are other colors in it".

Thus the pupils become familiar with quantifiers. Thus they will see that to classify objects each of which is of a single color only, there will be as many piles as there are colors, but in the case of objects having many colors, there will be as many piles possible as two raised to the power which is the number of colors, minus one. For example, taking green, red and red-green leaves, the two colors give us

$$2^2 - 1 = 3 \text{ possible piles}$$

But, taking the set of 3 colours

$$\left\{\text{red, yellow, green}\right\},$$

the following types of leaf are possible:

red, yellow, green, red-yellow, red-green, yellow-green, red-yellow-green where

$$2^3 - 1 = 7 \text{ possible piles}$$

In the case of leaves of the same color, there are no common elements between the piles of a color. In the more general case, there are common elements possible, between the set of leaves with red on, and the set of leaves with yellow on. "The whole leaf is red" excludes the possibility of another color in the leaf, whereas "there is some red in the leaf" does not at all imply such an exclusion. The experiment done with the autumn leaves can therefore serve as an introduction to the important properties of the logical quantifiers.

It is not necessary to use complicated terms. Children use the words "all" and "there is" and they realize that these words, along with words expressing properties, serve to identify properties of certain objects.

Leaves could also be categorized according to their shapes. Take for example the maple leaves, which have a different shape from other leaves. The decision as to whether a leaf is large or small will direct the children's attention toward ordering tasks. It is easier to determine whether a leaf is "larger than" or "smaller than" or "as large as" another leaf, than to say whether it is simply small or large. The pupils will also think of the relation "redder", "not as red", "more colors in it" or "less colors in it" after their work on classification according to color. It will be necessary to decide whether or not certain irregularities "count"! The children will learn to single out the properties for consideration and will not take into account other properties irrelevant to the categorization they are preparing. Thus, it can be seen that logical and relational situations can be introduced and discussed profitably, by introducing simple and commonplace material from the children's own environment.

1.6 Mathematics laboratory. Structured materials.

It is undeniable that all the mathematical relations to be learned by the children are too numerous and complex to allow us to confine ourselves to material taken exclusively from the immediate environment. For example, it seems that complex mathematical operations such as: multiplications of a multiple digit number by another multiple digit number can never occur in the physical environment. We do not mean that the circumstances in which one number must be multiplied by another do not occur, but we never in fact come across situations in which the actions involved in what is done mentally in such a multiplication are carried out concretely with real objects. For example, when we multiply twelve by thirteen, we think of twelve as being compounded of ten and of two. When we multiply by thirteen, we multiply by ten and we multiply by three. We take three times two and three times ten, then ten times two and ten times ten, and finally we add all these products to one another. In practice to realize these mental operations, twelve objects are needed; they must first be split into two sets, one of which would contain ten elements and the other two. For each element of these sets another set must be constructed which would contain thirteen elements. Each of these sets of thirteen should be subdivided into one set of and one set of three. Each time we find a sufficient number of objects to make more sets of ten, this must be done. When enough sets of ten have been found to put them into sets of a hundred, it must also be done. This kind of activity of grouping and regrouping sets into more numerous sets, can never be found in the physical environment. Consequently, if we wish the child to be intrinsicly motivated through being in personal contact with his surroundings these surroundings ought to be suitably enriched.

The concrete materials will serve to create similar activities made up specially so that the child is able to put into concrete form the operations which he is not yet capable of carrying out mentally. Such material is for example the multi-base material, the polymath material or the Cuisenaire rods. All this material is so thought out that the properties of the objects introduced are clearly and precisely observable; it is not so in the case of the autumn leaves, of the shells or of the match boxes. Thus in using these objects, we can reproduce, on a practical level, abstract mathematical situations that the child, in general, would not understand because he has never been in a position to observe personally, on the concrete level, the "events" described by the corresponding mathematical symbolism. For that reason, material must also be introduced for the learning of logic and relations; a set of logic blocks will then be used as well as a people set, trimath, quadrimath and so on, where each relevant property is well specified. Each property is or is not possessed by a given element of such a set. In the following chapter we shall see how it is possible to set up such a laboratory, consisting of an environment artificially constructed by means of structured and of "naturally" occurring material of all sorts.

1.7 The social factor.

We have discussed the role of the "physical" environment in learning, that is, in the adaptation process of the child in relation to his surroundings. It must not be forgotten that the social sphere is a very important part of that environment. Every child plays his own part; this fact leads to intellectual and affective communication between children. The situation should not be any different in a classroom: The teacher and the children create a social situation, be it an artificial one, and the latter leads to social contacts which, between the children, are made through play: that is his very important way of adaptation.

Social games can also be used to encourage and plan different kinds of learning. The teacher must also turn his attention to the adaptation process itself. If an adult is not adapted to his social circle, he is considered maladjusted.

Children play

a) by means of verbal communications,
b) by means of physical communication.

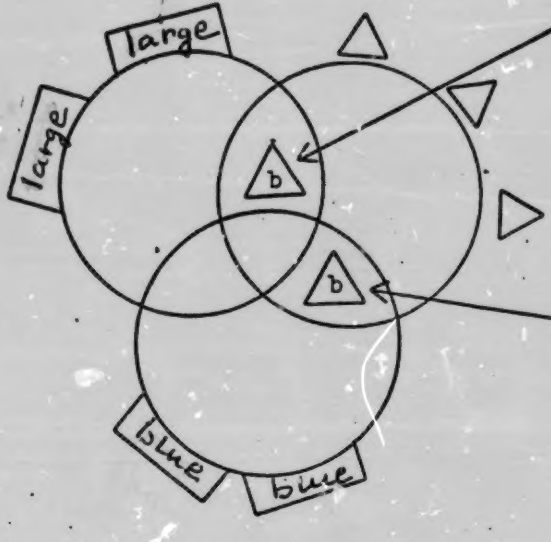
Verbal communication is carried on through the children's mother tongue. Physical communication is achieved through other sense modalities such as sight, hearing, smell, touch and the kinaesthetic sense. Children touch each other, look at each other and move in an organised manner when they play. They make rules for their play and these rules place them in a situation where their actions can interfere with one another. In this manner, they adapt to each other. They learn that they frequently have to restrain their desires to have the satisfaction of sharing the work or play of others; if each does as he wishes, it is impossible to play together.

The teacher is well advised to use social motiviation, namely the desire of a child to be with other children, in work or play. In traditional conditions, children have had to keep quiet. They have been forbidden to communicate among themselves and rarely been able to do so with their teacher, and even then exchanges have only been possible for one child at a time. Consequently, in the traditional schoolroom communication in general does not form part of the learning process. The organic situations of everyday life are always full of communications of all kinds. When we learn something in everyday life, we are surrounded by our social circle, by our "physical" environment, and the events produced in that milieu help us in our adaptation process. Thus it seems desirable to use the social factor in encouraging the learning process. Social contact is a factor which is very neglected with regard to methods used to motivate children. Children must keep quiet during a whole class; why not ask them to close their eyes? Maybe they would listen more attentively? Everyone will agree that this would be absurd. Interaction by sight is considered important in learning, why not social interaction?

The educational trick consists in giving a group of children an activity adapted to the social situation of the group: A game consisting of filling in a Venn diagram always leads to a lot of discussion. It is not obvious in the beginning, in which part of which hoop an object must be placed, even if the rules have been carefully set up. The application of the rules in a given situation is not always immediate. Consequently, discussions of this sort will be possible: One of the children might say: "No, this object should be placed here", another will disagree and an intellectual debate will follow, the decision will be taken according to the rules and the one who is right will win the debate; the winner will be able to demonstrate to the others that a certain object must in effect be placed in a certain spot, because otherwise the rules would not be observed. Thus the children will learn to discuss among themselves and acquire a healthy respect for the truth, the latter being the final authority in a discussion. Respect of the truth should have priority over the opinion we might have of someone in particular, such as that of the teacher. Even he can sometimes be wrong. The teacher's authority should never become absolute, it should anyhow be eventually replaced by the authority of the truth, which is itself absolute.

To use another example, a triangular network could be filled by placing three kinds of objects at each intersection of the lines. The rule will consist in never placing two objects of the same kind one beside the other. If two or three children begin the game by filling the network starting each from a different end of the network, it is very likely that the objects will not "fit", when they meet in the middle. The result will be an argument. It will be seen that something has to be changed. The players will decide if the change is to be made on one side or on the other. Generally, the children realize that it is not worth discussing whose pieces should be left standing as they know that whether they begin from one side or from another will not change the nature of the problem. What they must do is to satisfy the conditions imposed by the rules of the game and never place two similar objects one beside the other.

This is the kind of "truth-finding" conversation that we might hear:



John: I will place my block here because it is large and it is also △ , therefore it must go in the two hoops at the same time.

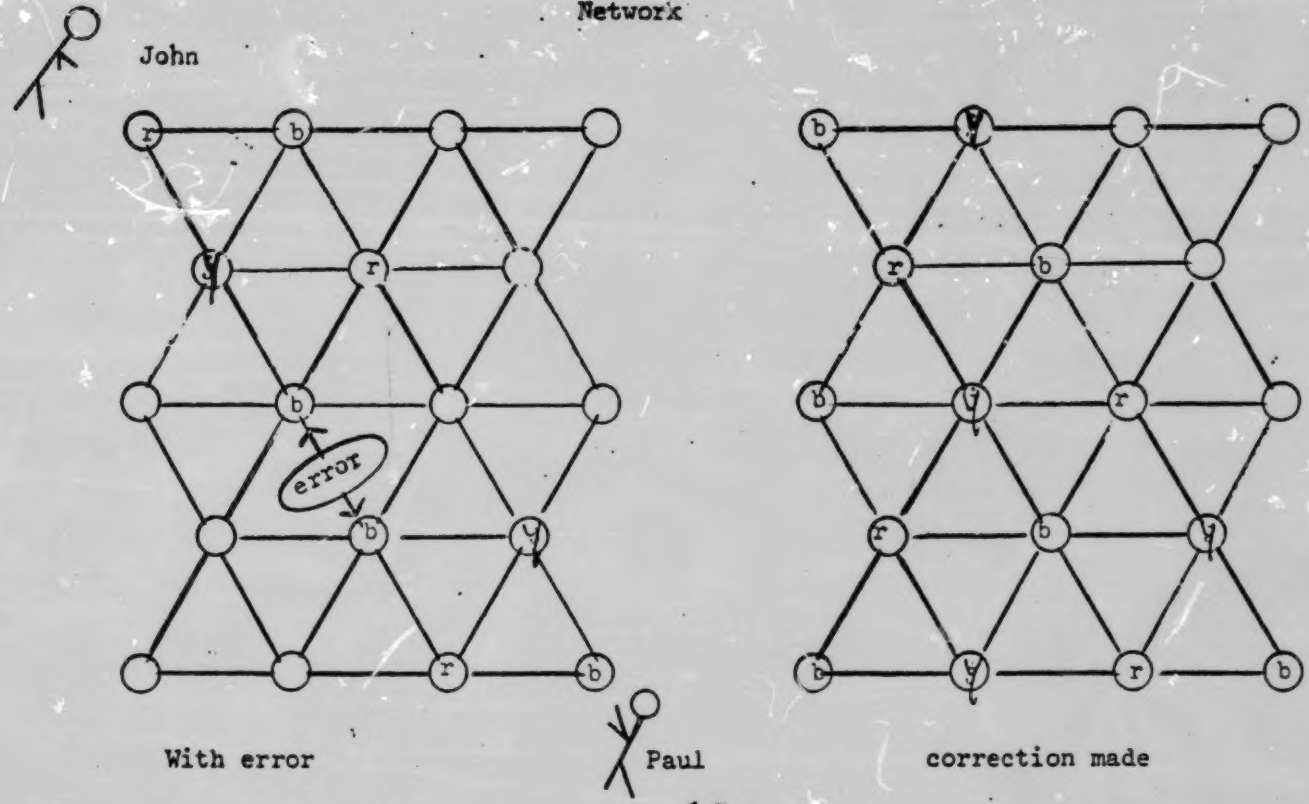
Paul: You are wrong, John, your △ is blue and you did not put it in the hoop of the blues!

John: Yes, I see, so I shall put it here. It is blue, therefore it will be with the other blue ones!

Paul: No, John, you have forgotten that it is large. It must go in the middle.

John: You are right.

Network



With error

correction made

1.8 Teamwork

The organisation of work in small groups raises a problem. Some teachers decide who shall be in a group. Others prefer to leave the children completely free to form their groups and choose their group leaders. Evidently there can be a happy medium. For example, a certain number of groups can be made up. If the then want to change their teams or their leader, their wish can be discussed and they can arrive at a compromise. It is sure that if the children can choose their companions intelligently the teams will last longer. Evidently they must get used to this. The exercise of free choice requires a certain amount of experience and children who are not used to it will find it more difficult than those who have already had to make many choices in their work. In general, smaller children can be grouped in teams of five or six; the older ones, from seven to eight years up, can be grouped in teams of four, and later on, in teams of two. A team of four is stable enough; such a team can be subdivided into pairs in three different ways, these pairs can work in teams of two and then, after having come to certain conclusions, the pairs can join together and discuss their problems. A team of three is unstable, a team of five is often too numerous to last long.

We have talked of social interaction between children. The interaction between the teacher and the children must not be neglected. The teacher will become the adviser, he will go around the classroom, giving advice and help to the teams who ask for it. A remark or two is sometimes sufficient to stimulate a discussion lasting several minutes if the remark is discerning and if it vitally concerns the particular situation in which the group finds itself. The teacher will call on his own judgment in dealing with the teams of children he supervises. After a year, certainly the teams will work more efficiently than on the first day of their formation. Evidently, in the Schools of Education, as well as in the Faculties of Education of the Universities this sort of practice should be taught, but in practical ways and not by means of lectures.

There will be occasions when the teacher will want to assemble the whole class and discuss certain problems with all the children. In mathematics, collective discussions with the whole class will occur less and less often, because the teams, and the individual children, will progress at different rates. Many will progress rapidly, others more slowly. This is in the nature of things. We are not all made the same and we must not expect to do the same things in the same period of time; to force children of one class to do so would be fruitless and not much good could come of it. The industrial world, for example, could not survive for one day under such conditions; it must be efficient because of the competition. In the educational field, there is no competition; all children must go to school and most of them go to public schools; consequently, even any real competition provided by private schools is missing. The latter, for the most part slavishly imitate what the public schools do. For this, and many other reasons, the efficiency of learning in a traditional classroom is terribly low.

Learning efficiency can easily be increased by a very high percentage. For example, in the Sherbrooke project, arithmetical techniques are taught during one fifth of the total time. These same techniques form for the most part the entire curriculum, in traditional teaching. The Sherbrooke project schools nevertheless attain approximately the same level of arithmetical competence by grade four with only twenty per cent of the effort directed towards arithmetic. This means that according to this rough measure, learning efficiency has been increased 500 per cent. Children are taught the same techniques in one-fifth of the time. not to mention the fact that during the remainder of the time, the children in the Sherbrooke project learn mathematical notions quite unknown to those in the traditional system.

A large part of this increase in "efficiency" is probably due to the child's freedom of action in class, that is, to the fact that he is not forced to progress at the same rythm as the least gifted child in the class. The less gifted children need more material - which must be more varied and they need much more practice. The more gifted ones, on the other hand can generalize and abstract more rapidly. They need less material and less practice, and can consequently move on from one mathematical idea to another much more rapidly. This situation need not be deplored, variability being the essence of all organic and living situations. If the less gifted are put in one class and the more gifted ones in another, the situation will deteriorate because the gifted children will compete with one another whereas in the classes of less intelligent children, there will be less intellectual stimulus provided by a greater variety of activities engaged in by the more gifted.

The separation of children into homogeneous classes is a pedagogical crime of the first magnitude. Once the child has become aware of his position in the hierarchy set up by the school, there he will remain, probably during the whole of his career. The school will expect the child in a "B class" to behave like a "B child". It is clear that such a classification of the child is a sentence which condemns him at a very early age to an intellectual fate that he certainly has not deserved. In two or three years, it is quite possible that a child, slightly retarded in relation to the most brilliant, will make considerable progress, if suitably stimulated. He will clearly not do so if he is put in his special class where he is not expected to make such progress.

The elimination of lecturing at children and of homogeneous classes appear to be an urgent duty of modern pedagogy. An intellectually rich social environment is more likely to encourage than to discourage the child in his progress, in the intellectual as well as in the social and affective spheres.

1.9 The role of tradition.

A child at birth finds himself in a situation whose nature is determined by historical, economic and social forces. Some of these forces tend to conserve, to perpetuate the existing situation and others tend to bring about changes. The behaviour of an individual at any given moment in history, at any given place in the world, is to a large extent determined by conservative forces, called tradition. Tha values and habits manifested by such a person are largely conditioned by the traditions of that period, and of the particular place in question. One need only look at the history of humanity to realize to what extent traditions have varied.

Traditions evidently influence to a great extent the behaviour of individuals who live in the midst of any particular set of traditions. If we wish to understand what the learning process is about, in particular the process of learning mathematics, an effort must be made to understand the role played by tradition in that learning process. The attitude of the child in such a situation is conditioned by traditional sets of values and habits that he encounters in his surroundings. The relative calm found in the classrooms of Quebec in relation to those of the United States could very well be a consequence of the relative calm at the social and political levels in Quebec and in Canada in general, in contrast to the social disorders prevelant in the U.S.A. This does not mean that Quebec and the rest of Canada have no economic problems; but if we compare the intensities of the internal conflicts of both countries, it seems to the casual observer there is a tradition of rebellion in the United States whereas in Canada, and particularly in Quebec, rebellious behaviour is much more recent and has not yet had a great deal of influence on children. American visitors are astonished at the calm and the facility with which Sherbrooke teachers can handle a class of thirty children. How is it that the children keep on with their work in the teacher's absence or without his immediate supervision? According to the opinion of numerous educationalists familiar with both situations the American teacher has to have his eyes open and know what is going on at every moment, in the classroom; otherwise, he quickly loses control.

This difference could possibly be due to the pervasive but subtle influence of the Catholic Church in the province of Quebec. Only a few years ago, the authority of the Church was almost absolute. Even today, children are subjected to a spiritual discipline linked to a mental and physical discipline. Thus, a class can more easily be organized in teams in Quebec than in the U.S. Take the situation in New York, as an extreme case on the U.S. side, where a positive attitude towards discipline is definitely not part of the local tradition. On the contrary, New York local tradition, especially in the ghettos, encourages an attitude of opposition, of rebellion on the part of children against the teacher's authority, in contrast to the general attitude of respect for the adult shown by Quebec children.

The above remarks are not intended as moral judgment. Respect of the teacher's authority, as has been seen, is not always an absolute advantage, even though a total lack of respect for the teacher's personal authority could naturally have disastrous consequences.

In short, methods of classroom organization must take into account the local cultural trends to which the children are necessarily subjected.

Let us now consider certain less fundamental traditions. Take the use of grades. Once the child has realized that the school authorities give a high grade as a reward for work well done, he will not want to work without this reward. The child's involvement in the work done is no longer his motivation. If the child remains in this situation of extrinsic rewards long engouh, he will continue to expect such a reward because it is part of the set of traditions to which he has become accustomed. The adaptation process forces him to get used to it and finally, he finds the situation provided by the traditions indispensable to continued activity. This result is of course the inevitable consequences of working within the narrow limits of any system of traditions.

Take another example: In general, the children in North America must become adults at an early age. To play with toys is practically unthinkable for a child of 12: dolls have to be put away, trains, cars and other toys of this kind are abandoned. Children play hockey, they ski, they fight, but they are not interested in toys; consequently, if concrete material is presented them at an age at which, according to tradition, "toys" are already things of the past, the children will resist the introduction of such material, taking it for "Kidstuff"; for them, to be a child is a disadvantage. On the other hand, if in a school or in a whole school system, a tradition is established that concrete learning material is an integral part of learning, there will be no such objections. Children will realize that in all the classes, concrete material is used and that it helps to solve problems. This new tradition, in the particular subculture in which it is established, will influence the child in place of the general tradition which tends to forbid games at a very early age in the North-American culture.

1.10 Evolution

Certain traditions are established: they are the logical consequences of a century-old economic and cultural development. Among these traditions, we find the one of compulsary attendance at school, which was introduced in most western countries about a century ago. It was necessary after the industrial revolution, that the population in general be taught reading, writing and arithmetic, so that industries could function efficiently. Consequently, such basic learning soon became the only aim of public education in most countries. What is more, this tradition developed very fast. It became apparent that children had to meet these minimum requirements as early as possible, so they could be absorbed by a growing labour market. If a child did not meet these requirements, he was considered retarded in relation to the others and, once out of school, his chances of success in life were small. That tradition has remained with us to this day in spite of the fact that the situation has completely changed. Children go to school till the age of 16 or 17 and even 18, depending on the country. However, certain teachers and parents worry - without reason - when they see a child of 8 who can not add, subtract and multiply; this is due to the fact that they do not know, and are on the whole not in a position to know what mathematical skills are important today. These are radically different from those of the last century. Certain traditions subsist and are accepted by the population in general, the latter not questioning the why and the wherefore of their existence; so it is very difficult to change the situation in terms of the requirements of the modern world. It seems that not enough information is flowing from the source to the public as to what are important skills today for those who wish to make progress within our existing social system.

The mathematical skills required of children today should not be the same as those required of the children of yesterday who had to obtain a job at a factory at a young age a century ago. At every parent-teacher meeting the ritual question is still asked: "How come my child does not know 2 and 3 make 5 whereas at 6, I knew my multiplications? "These parents are unaware of the reasons for which it was necessary some time ago to acquire mechanical skills in arithmetic. One of the reasons was that most of the children left school at an earlier age; moreover, the requirements were different from what they are to-day.

We now have computers to do all sorts of calculations; it is becoming less and less important to know to perfection the so-called tables which were once so useful. What we need is to know when and under what conditions to use which kind of operation. The child in the traditional school often asks when confronted with a mathematical problem, if he must add or subtract, multiply or divide! Even if he can perform these operations, he does not know when to apply them; this means that he becomes a machine without an operator because he does not know which operation to apply and when. Today it is very important that children should know the reason for using any given operation and that they should be aware of the internal links between the mathematical operations they learn; it matters less that they become mediocre little "calculators". The new tradition, in favor of such a concept is slowly developing; this is what we might call: "evolution". Public opinion, conscious of the requirements of modern living, is slowly changing. Mechanical work is performed much better by machines; on the other hand understanding, coupled with ability to manipulate these machines, becomes of prime importance.

In the same way the old traditions must gradually change with regard to class organization. About ten years ago, in certain countries, Canada included, no one would have thought of organizing a classroom in any other way than in rows, with the teacher in front pontificating. Today the thought of a decentralized class, with a corner for reading, another for painting, another for science, another for mathematical material, and so on... is accepted by a larger section of the public. Such experiments in initiating changes are favoured by school boards whose members can discern the totally different conditions of learning to-day, if such learning is to become increasingly efficient.

As has already been mentioned, learning efficiency ought to be increased enormously. One way to ensure that this will happen is to change the old established form of classroom organization; in short, we should abolish homogeneous classes and introduce work in small groups. Public opinion, as well as the more specialized opinion of educationalists across the country, is beginning to change. The children in the classrooms are also subjected to these evolutionary tendencies. They note some definite changes in the teachers' conversation and in parental attitudes. They realize that the teacher is not always in a position to give them a precise answer. This might lead to a certain lack of security in the case of some children. However, they must adapt to this change as they must adapt to all changes in the environment in which they develop. Some teachers, recently trained, have been encouraged during their training to make deliberate mistakes so that the child loses confidence in absolute authority by being forced to ponder the answer and check for himself if the teacher is right or wrong. The public in general must be conscious of these evolutionary tendencies, so as to understand the situation of the teacher and of the child in a mathematics class in the Quebec of today.

1.11 Revolution

There are certain attitudes, certain forms of behaviour which can be changed gradually. However, there are others which can only be modified all at once. For example, subject matter radically different from those contained in the traditional programmes can only be introduced all at once. Logic is either taught or not taught. Sets are either taught or not taught. There is no half way house here! Knowledge cannot be "half-given". It is clear that in some cases, a decision must be taken as to whether a radical change is to be made or whether the "traditional" situation should remain. The same argument applies in the case of behaviour; either an attitude of authority is adopted in relation to the child, or the child is considered as a living being in his own right, growing and developing. We act in terms of one attitude or the other; it is impossible to keep an attitude of authority and at the same time to reject it for a more flexible one. Revolutionary changes have taken place in the child's environment as well as in the immediate conditions under which he works in the classroom. The introduction of a new subject to which the child is not accustomed is for him a revolutionary change.

The other day, in a school, I heard children singing "O Canada" and I noticed that they did not do so in a very convincing way; I said so to the teacher, who answered: "You see, they don't have their heart in it"! That means that those children had already been influenced by the revolutionary events going on in Quebec. They hear of a revolution, of the separation of Quebec from the rest of Canada, and when they have to sing "O Canada" they are ill at ease. The revolutionary situation observed in the province must permeate the life, the attitude and the behaviour of the children. Undoubtedly, the children find it unreasonable that in their surroundings, most people are against Federation and that as soon as they are in school, they must sing their national anthem. The children do not dare to say "What for?" because the changes have in effect not yet taken place. However, there is a change of attitude in the minds of those who have thought over the problem. These individuals have taken a decision, favoring one side or the other and children are influenced by these decisions and by the accompanying changes of attitude.

In this chapter, I have tried to explain how the child's environment must be taken in consideration in discussing the details of his mathematical training. The physical aspects of this environment, social factors and those resulting from the interaction of tradition, evolution and revolution, provide the cultural background, against which the child must find some kind of equilibrium. In all the discussions of the later chapters we shall take into account these fundamental conditions which determine the development of his attitudes, of his ways of behaving in relation to the educational reforms we would like to see put into practice in Quebec schools.

Chapter 2

Pre-mathematics

2.1 Discrimination

During the adaptation process of the organism to its environment, classifications must be made, which means that certain consequences of our actions help us to consider phenomena as equivalent. For example: if a mad dog comes near us, prudence forces us to run away or to defend ourselves if he attacks us, but we cannot keep calm. We notice the ferocity of this animal as well as some of its other characteristics, such as the length of its tail, the color of its fur... but these characteristics will not influence our decision. In other words, in such circumstances a certain number of dogs will be considered as similar while others will be different. A distinction will be made between ferocious and not ferocious.

Take a sum of money of for example, 25 cents. Whether it is a twenty-five cent piece or two ten-cent pieces and one five-cent piece, is of no importance to us. The distribution of the sum does not matter as long as the total is the same. The two events will seem to us to be equivalent, whereas a child who does not know the value of the pieces will prefer the set of three pieces. It must also be realized that what is similar in one case will not appear similar in another. For example, if we wish to go from one place to another by taxi, the color of the car is not important, a taxi of one color will be the same to us as a taxi of another color. But if we wish to buy a car, we will choose the color that we like best. However, if we see a car coming towards us at high speed, we will do everything possible to avoid a collision without looking at the color of the car. Evidently, one car will be equivalent to another, depending on the circumstances. The organism must use its common sense in order to survive. Thus he is forced to consider objects with a view to classifying them, so as to ignore the unessential differences between these objects in a particular situation. He must do so in order to ensure his survival and generally the satisfaction of his daily needs.

Discrimination between different objects and different events develops quite early in a child. At two years he is already capable of distinguishing shapes, and he can tell most of the colors at three. Later the individual learns to distinguish between much more complicated situations. This is what happens when a child begins to understand conceptual discrimination in mathematics. He must recognize certain similarities and certain differences between more and more complex situations. Thus he will be in a position to use his sense of discrimination on the abstract and general level in mathematics.

Material such as logic blocks, trimath, or quadrimath, serve to emphasize and encourage the process of learning to recognize similarities and differences and to make this process more conscious and less random. Suppose we have two blocks, with four attributes we consider relevant. The child has to decide if, as far as these four attributes are concerned, there is a difference or a similarity between the two blocks. For example: a large, red, thick square and a large blue, thick circle are similar in size and thickness but are different in color and shape. When the child is able to become aware of these differences and similarities explicitly, he will eventually be able to use them to put into concrete form mathematical structures in which similarities and differences of the kind are relevant. Suppose the child is already in a position to discern all the differences and similarities between two logic blocks and that he can state them explicitly; he could then be asked to distinguish the relations between one block and another and those between a third and a fourth block. Suppose, for example, that we take a large, blue, thick square and a large blue thin square; the thickness is the only difference between these two. The child can be asked to take any other block and to find a fourth block related to the third in the same way that the first was related to the second. Indeed, this type of discrimination between relations is commonly used in the construction of I.Q. tests.

Let us take a people set (a set of plastic figures), which has precisely the same difference-structure as a set of logic blocks. Let us show the child a blue walking woman and a blue sitting boy; the child must find the three differences, namely position, size and sex, and the one similarity, namely colour, between these two members of the people set. Starting with any other "person" in the people set , he should be able to form another pair with this person as one member, in such a way that the resemblances and differences between the members of this pair are the same as those between the members of the first pair. The game which consists of constructing all possible pairs is one which encourages the development of such distinctions at the conceptual level. In this type of situation, the child has to compare, not one object with another, but one relation with another. He must say to himself: "this relation resembles that other relation", or "this relation is different from that other relation."

2.2 Discrimination between orders

The same kind of discrimination learning can be called into play when dealing with order relations. Take six objects of three different colors and two different shapes and arrange them in a certain way and then in another way We would really need two sets of the same kinds of objects so that the child can look at both arrangements at the same time. For example, on the left we have an arrangement where all the squares come before all the circles and a red one comes before a blue one, a blue one before a yellow one, for blocks of the same shape; that is, there is red square, blue square, yellow square, red circle, blue circle, yellow circle. In the second arrangement we have a red square, a red circle, a blue square, a blue circle, a yellow square, and a yellow circle. The child will immediately see that the way in which the first set is arranged does not resemble the way the second one is arranged; the difference is that in the first case all the squares come before all the circles whereas in the second, all the red ones come before all the blue ones and all the blue ones before the yellow ones. How could they be made similar?, Two sets could be placed in precisely the same order, and then the child could see for himself that the first arrangement is identical to the second one. But if only the order of the colors were changed, the second arrangement would be:

yellow square, blue square, red square,
yellow circle, blue circle, red circle.

The method of arranging the six objects first according to shape and then according to color, will have been kept in both arrangements. The difference in both cases exists only in the way the colors are placed one after the other. Here, by "resemblance" we understand that priority is given to shape and that the colors follow one another in the same manner for each of the two shapes within the same arrangement of blocks, we now say that these arrangements are "the same". The circles could have been placed before the squares and the changed arrangement would still have remained the same! For example; yellow circle, blue circle, red circle, yellow square, blue square, red square, could have been chosen. Here at first the shapes were arranged in two parts, the circles on one side and the squares on the other, then they were arranged according to their color, but in the same way for both shapes; the different color circles follow one another in the same succession as the different color squares. Thus from this point of view it can be said that both orders are "the same".

In short , to compare the orders of arrangements one can consider:

i) the priority of one criterion over another,

ii) whether the second criterion is applied using the same order for each set resulting from applying the first criterion,

iii) whether the order of elements is partially or totally identical in the two arrangements.

In our examples so far we have considered (i) and (ii). We have said that two arrangements follow "the same" order if conditions i) and ii) are satisfied, not necessarily condition iii). This means that to establish an order which "resembles" our first arrangement using criteria (i) and (ii), the order of shapes must have priority over the order of colours and the order in which the colors follow one another must be the same for the two different shapes.

Obviously a stricter criterion could have been applied: to be "the same" could mean simply that the elements must follow one another in exactly the same way in both arrangements. Or we could relax our conditions for "sameness" and say simply that two orders are "the same" if the order of priority of the different attributes is the same in both arrangements. According to the relaxed criterion the following orders:

red square, blue square, yellow square,
red circle, blue circle, yellow circle,

are similar to

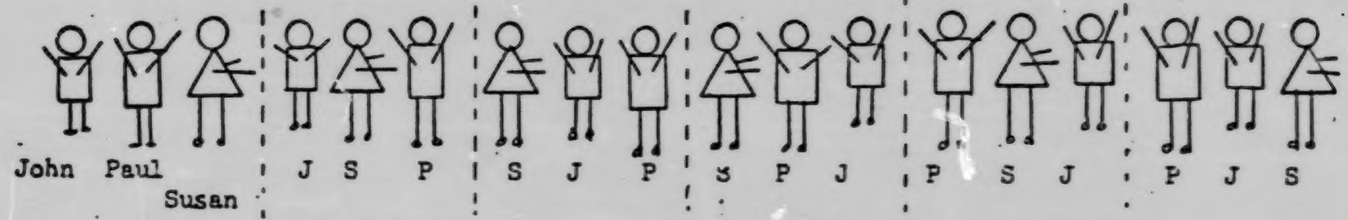
blue circle, red circle, yellow circle,
yellow square, red square, blue square.

In both cases shape has priority over color, and although the order of the colors in each of these two arrangements is not the same we are not counting this as being "different"!

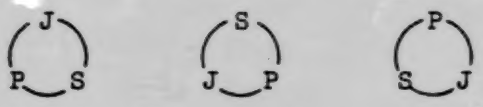
The child learns to consider resemblances and differences according to criteria previously determined. One thing resembles another if, in accordance with a previous decision, the criteria of resemblance are satisfied; otherwise the arrangements are, by definition, different.

For another example we could take permutations of different orders

Take three children, John, Paul and Susan, and put them in a row.

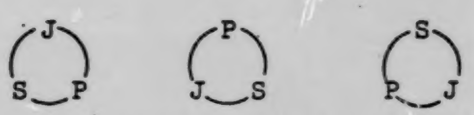


There are six possible ways of getting them to sit on the chairs provided, these being placed in a row. If on the other hand, the children are asked to hold hands and make a ring, there will only be two different ways of doing this if we do not count where each child is standing but only whose hand he is holding in his right hand and whose hand he is holding in his left hand.



will therefore count as one and the same arrangement

and



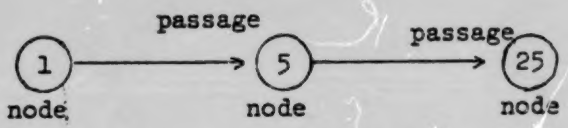
will all count as just one arrangement

and there is no other way.

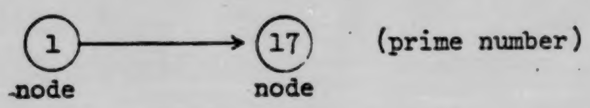
Thus John, Paul, Susan and Susan, John, Paul; Paul, Susan, John, here, will be similar arrangements because these three arrangements result in the same kind of ring, whereas if the children are sitting on chairs in a row the corresponding three arrangements will appear to be quite different.

2.3 Networks

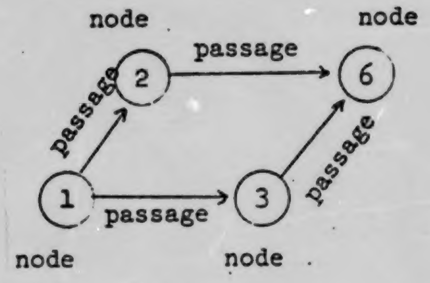
Other examples may be found in other parts of mathematics. Consider prime factors. If a network is made of these factors starting with 1 and ending with the numbers in question,a diagram will result for each number; if a prime number is taken, there will only be 2 nodes in the network and only one passage. The square of a prime number will always have a network with 3 nodes and 2 passages; whereas a number such as 6 would have 4 nodes and 4 passages because starting from 1, we can multiply by 2, after 2 we continue by pultiplying by 3, and so we get to 6. We could have multiplied 1 by 3, then 3 by 2 to get to 6; thus we obtain 4 passages and four nodes.


passage
passage
1
5
25
node
node
node

square of a prime number


1
17
(prime number)
node
node

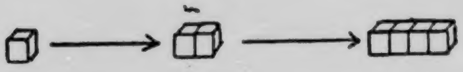
2 nodes and 1 passage


node
node
passage
passage
2
6
passage
1
3
passage
passage
node
node

4 nodes and
4 passages.

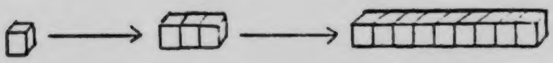
We will see that the number 4 will have the same type of network as the number 9. The number 6 will have the same type of network as the number 10. It is not obvious at the beginning that there is a structural ressemblance between the number 6, the number 10 and the number 14, and that these numbers, on the other hand, possess a structure different from 4 and 9 while these last two are similar from the point of view of the factorial structure.

4



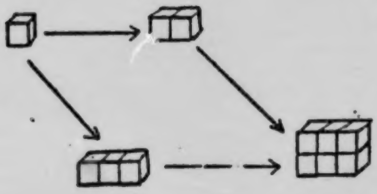
3 nodes
2 passages

9



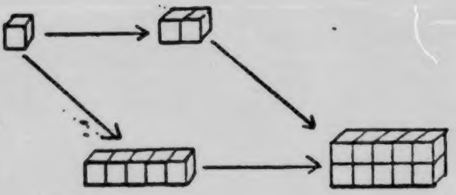
3 nodes
2 passages

6



4 nodes
4 passages

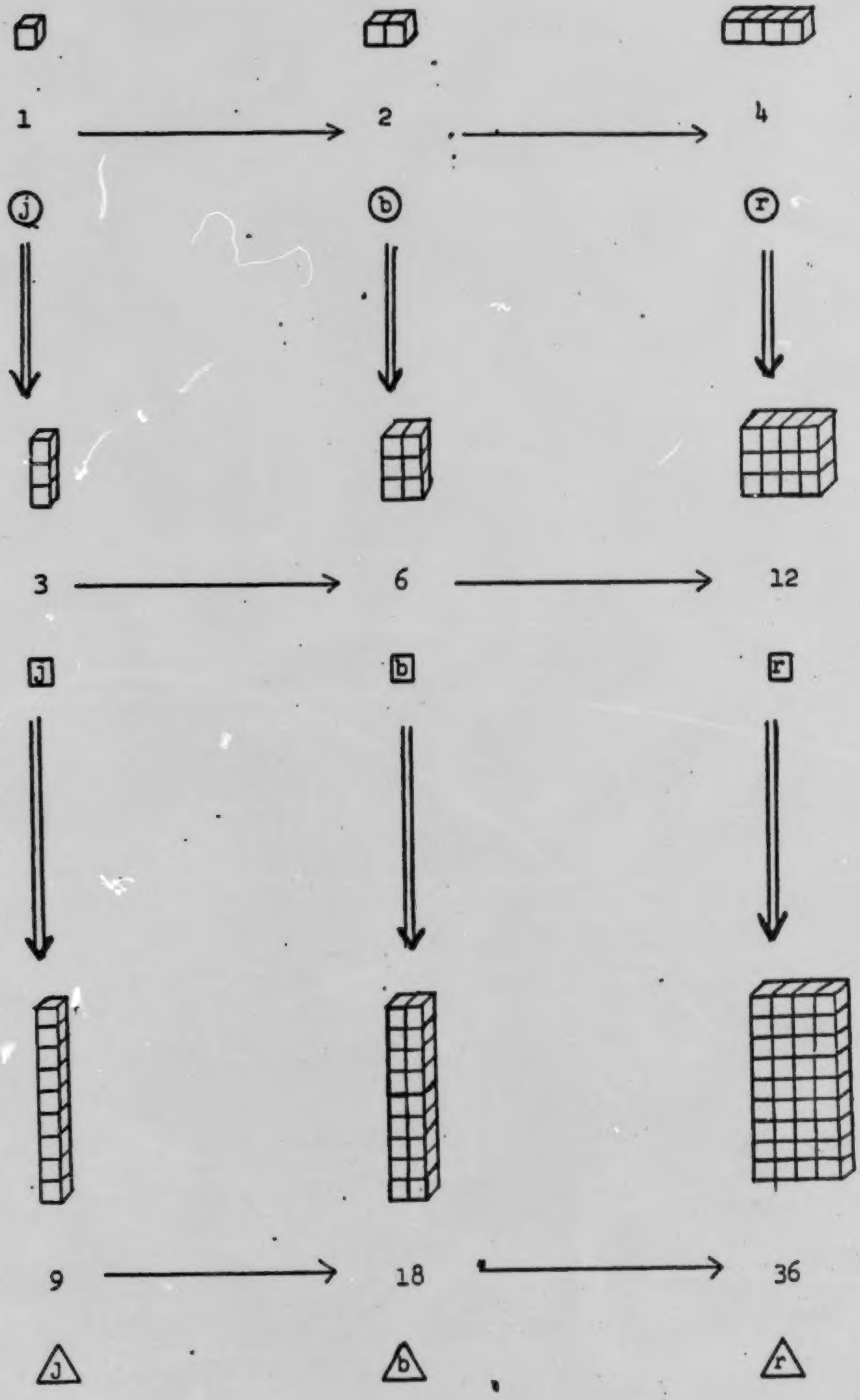
10



4 nodes
4 passages

Thus, a distinction on the mathematical level can be made between numbers by considering their factorial structure. Two numbers having the same factorial structure will be called similar and two numbers having different factorial structures will be called different.

Resemblances and differences, when it comes to somewhat more complex structures, can be described by means of arrow diagrams. An arrow diagram consists of nodes and arrows. In the last examples it was seen that the arrows represented multiplications by a prime number; to pass in the opposite direction of an arrow represented a division by a prime number. We could for example play with logic blocks having the same structures as those factor games; these more concrete games should preferably precede the games with factors. To understand the meaning of a factor, the nature of multiplication must be understood, and this is not understood as well as is generally admitted. For example, the factors of 36 could be represented by nine blocks, circles, squares, and triangles; red ones, blue ones, yellow ones. A yellow circle could represent number 1; a change from yellow to blue could represent a multiplication by 2 from blue to red also a multiplication by 2 while a change from red to yellow would be a division by 4 because we must go back to one by taking the two arrows in the opposite direction. Multiplication by 3 could be represented by going from a circle to a square or from a square to a triangle but of the same color. Therefore if the yellow circle is 1, the yellow square will be 3 and the yellow triangle will be 9. Thus we have the following diagram for the factors of 36 with the 9 logic blocks acting as a helping agent to emphasize the structure being learned.

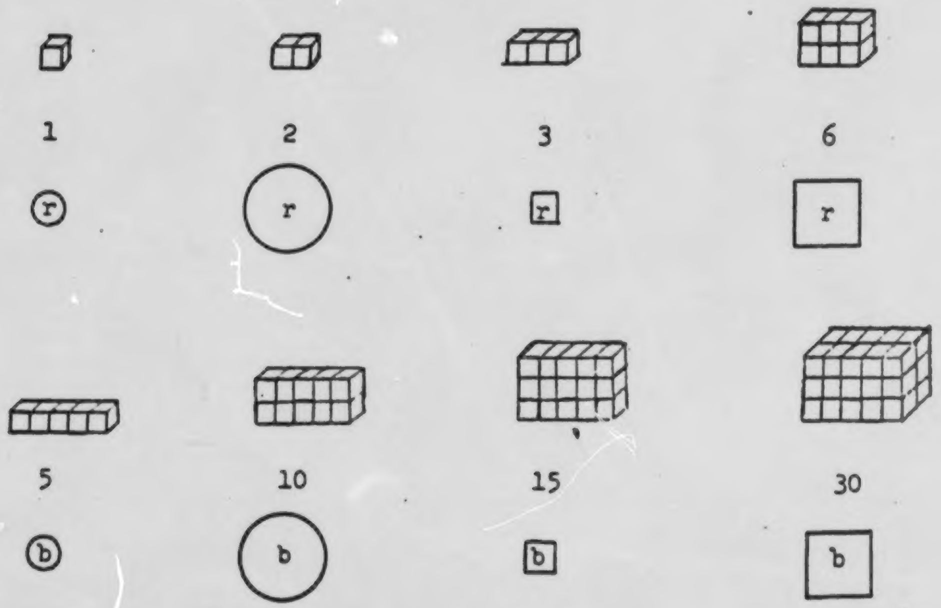

1
2
4
j
b
r
3
6
12
J
b
r
9
18
36
j
b
r

9 nodes
12 passages

The 9 possibilities in the factor game or in the logic block game can also be represented by sentences or by words. If sentences are taken, three subjects and three verbs could be chosen. For example: father, mother and Linda would be the subjects; is working , is eating and is sleeping could be the verbs. Thus nine sentences can be constructed because with each of the subjects one or the other of the three verbs can be associated. Each factor of 36 will have a sentence to which it corresponds. Transcoding games could be played -- they will be mentioned again later -- between the factors, the logic blocks and the sentences.

1. father is working
2. father is eating
4. father is sleeping
3. mother is working
6. mother is eating
12. mother is sleeping
9. Linda is working
18. Linda is eating
36. Linda is sleeping.

Other games can be played with eight blocks; for example: the large red circle, the small red circle, the large red square, and the small red square, the large blue circle and the small blue circle, the large blue square and the small blue square. An association could be made between this set of blocks and all the factors of 30; there are exactly eight: 1, 2, 3, 5, 6, 10, 15, and 30.



If sentences are used, two subjects, two verbs and two objects could be introduced; for example: "father" and "mother", is working and "is eating", in the "house" and "outside". Thus the following sentences will result, associated to the eight logic blocks and to the eight factors of 30.

1 father is working in the house ● (in red)
2 mother is working in the house • (in red)

3 father is eating in the house ■ (in red)
6 mother is eating in the house ▪ (in red)

5 father is working outside ● (in blue)
10 mother is working outside • (in blue)

15 father is eating outside ■ (in blue)
30 mother is eating outside ▪ (in blue)

Choose one pair of blocks, for example: the large red square and the small blue circle. The task consists in finding the other pairs in which there are just the same differences between the two members of the pair. Also, arrange factors of 30 in pairs, using the factorial properties of the numbers which must differ from one member of the pair to the other in the same way as this is the case for the first chosen pair. Take the number 6 which can be divided by 2 and by 3, but not by 5. To find the other member of the pair, three differences are needed. It is necessary that the other member of the pair not be divisible by 2, nor by 3, but by 5; consequently the number needed is 5; thus 6 and 5 belong to the same pair. Now let us look for the number that "goes with" 2. It will be divisible by 3 and by 5 but not by 2, and so the number must be 15. It will be noticed that the product of members of the same pair is always 30. In the case of sentences we shall have, for example:

father is working outside
which will be put with the sentence
mother is eating in the house
because everything must be changed, the subject, the verb and the complement. Thus, exercises involving similarities and differences

should be provided, in situations where the properties to be distinguished are well known to the children. Thus they can concentrate on perfecting their conceptual discrimination rather than be involved in unusual physical discriminations. For example, if children know about factors they can be asked: "If 5 is related to 15, 2 is related to what?"

The same relation must hold in the case of each pair, the pair (5, 15) and the pair (2, ?). What is the pair (2, ?) which is similar to the pair (5, 15)? With regard to possession or non-possession of a factor of 5, the two members of the pair must be similar because 5 and 15 both possess the factor 5. Number 2 does not possess factor 5, therefore the other member of the pair should not do so either. Consider the factor 2. This number is not a factor of 5 nor of 15 consequently as far as the "property 2" is concerned, 5 and 15 are similar. Consequently, since 2 has a factor 2, the other member of the pair must also have it as a factor so that the property of having or not having a factor 2 will be the same in both pairs. Consider factor 3: 5 has no factor 3 but 15 does; 2 has no factor 3 consequently the other member of the pair should have it.

The number we are looking for does not have the factor 5, does have the factor 3 and does have the factor 2; thus the second member of the pair is 6. It is evident that 5 is to 15 as 2 is to 6. The relation between 5 and 15, numerically, is also similar to the relation between 2 and 6. The numbers differ in a factor 3 in both pairs.

There are interesting links between possession and non-possession of certain factors, by numbers and relations involving color, shapes, thickness and size. In these differences and resemblances are determined by having the same color, or size, or shape, or thickness, or having differences in these attributes.

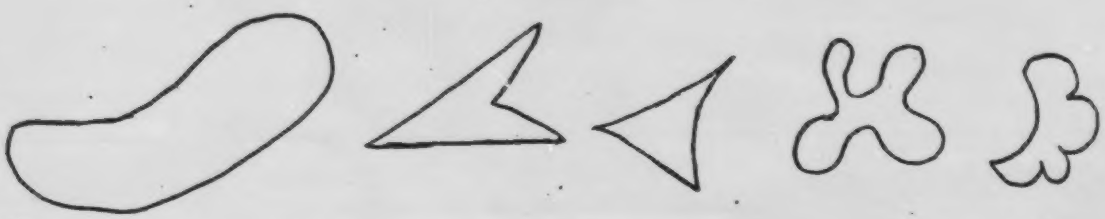
Games played with factors are not a part of pre-mathematics. The links between certain sets of factors, sets of sentences and certain sets of objects have been mentioned so as to underline the fact that structured games in pre-mathematics lead directly to important structures which the child will learn later on. It should be clear by now that such factorial structures can be introduced in principle through the use of similarity and difference relations, using objects, sentences or words. As soon as the child is able to understand the links between these different embodiments, he will have made his first steps towards real mathematics.

2.4 Various examples

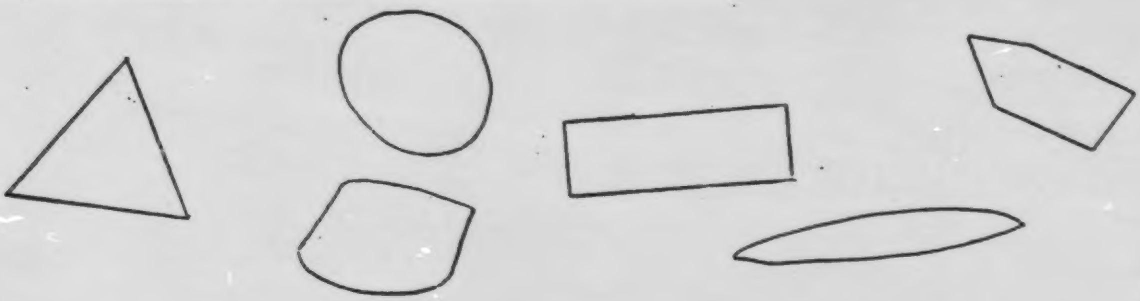
Pre-mathematical activities can be prepared in order to begin the learning of any mathematical topic.

To learn to classify, exercises aiming at the identification of certain properties, are necessary. Such properties must be specified. This is a static situation. Here is an example of a game of classification:

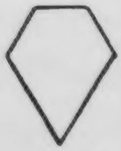
Here are some TRINGES:



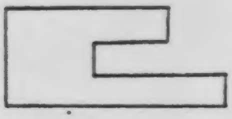
Here are some figures which are not TRINGES:



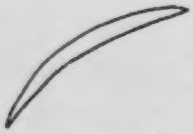
Which of these figures are TRINGES ?



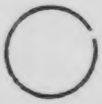
a



b



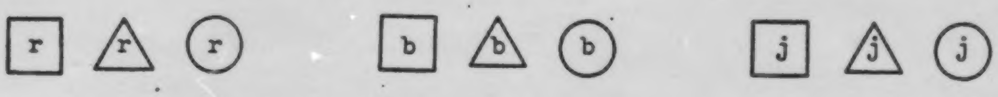
c



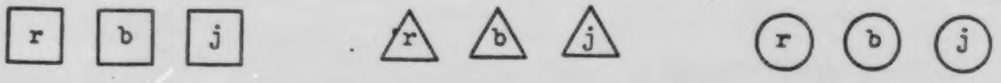
d

Answer: b and c

The objective must be that the child should realize the way in which the objects are related one to the other, starting from certain properties considered as relevant. There are various types of relations with which the budding mathematician will have to become familiar. The relation based on the identification of elements on account of a certain common property possessed by all elements so related, leads to equivalence relations. An element is said to resemble or to "be similar to" another element if both elements have a certain property in common. If the elements taken in consideration can be naturally subdivided into a certain number of disjoint sets of elements, in such a way that the similar elements always belong to the same set, and that no element belongs to more than one of these sets at the same time, then such a "similarity" relation is said to be an equivalence relation, for example:



is a subdivision of a starting set of nine blocks, where a similarity as regards color has been considered. Again, the subdivision



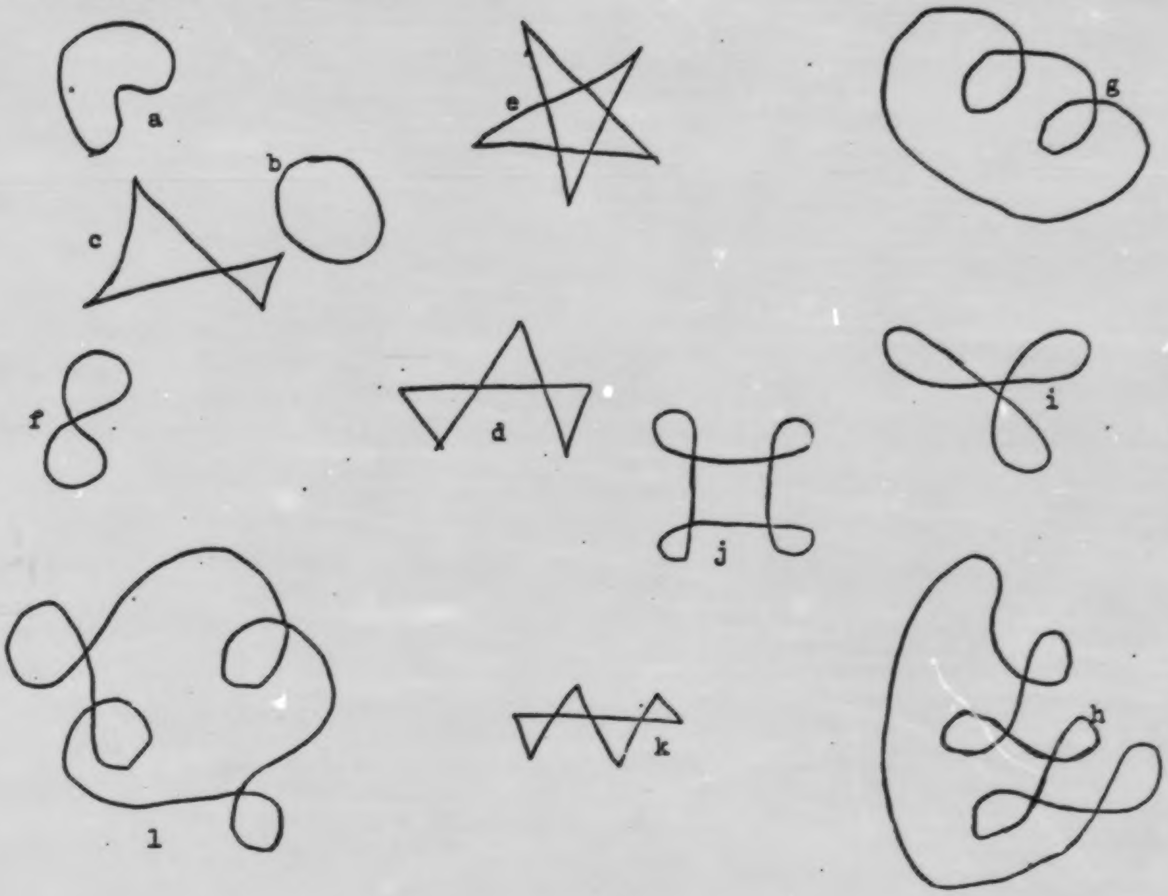
takes account of shape only. All the blocks of the same shape are considered "equivalent".

Possession of one shape excludes possession of another shape. For this reason there is no overlapping between the sets so constructed. In the case of our first subdivision, it is also evident that each block has only one color, and consequently "..... is the same color as ....." is an equivalence relation.

The fixed nature of classification is reminiscent of the role of tradition in society. Relations of order, fundamental in understnading numbers, may in some sense remind us of the process of evolution in society. One event leads to another according to laws which are very complex but nevertheless can be formulated by means of probabilistic functions.

This process of evolution is possibly the ultimate source, within our own personal and social experience, of the mathematical idea of order.

Games in which the child manipulates the ideas of "before", "after" and "at the same time" must precede the comprehension in depth of the mathematical idea of order. The following figures provide an example:



Suppose we know that:

ⓐ eats ⓒ, ⓕ eats ⓓ, ⓘ eats ⓛ, and ⓖ eats ⓘ.

The problem is one of determining other pairs of figures where the first one eats the second one.

Answers: ⓑ eats ⓒ, ⓑ eats ⓕ, ⓐ eats ⓕ, ⓛ eats ⓔ,
ⓘ eats ⓙ, ⓒ eats ⓓ, etc.

The figures could be related by using other criteria. For example:

ⓐ likes ⓑ, ⓔ likes ⓗ, ⓚ likes ⓚ, ⓗ likes ⓔ, ⓚ likes ⓘ

Then other pairs of figures must be found in which the first likes the second.

Answers: ⓛ likes ⓙ, ⓒ likes ⓕ, ⓓ likes ⓖ, etc. .

When one realizes a difference between two elements, such as, for example between two objects or two events, one does so from the fact that these two elements do not "belong" together, but "belong" to two different classes. A difference is always more striking than a resemblance. A radical change is noticed, when passing from one element to a different element. The abstract notion of a "difference relation" is formed by experiencing radical changes. Thus the understanding of changes of a sudden nature in our surrounding circumstances, helps us understand the properties of differences.

Here is a pre-mathematical exercise based on differences:

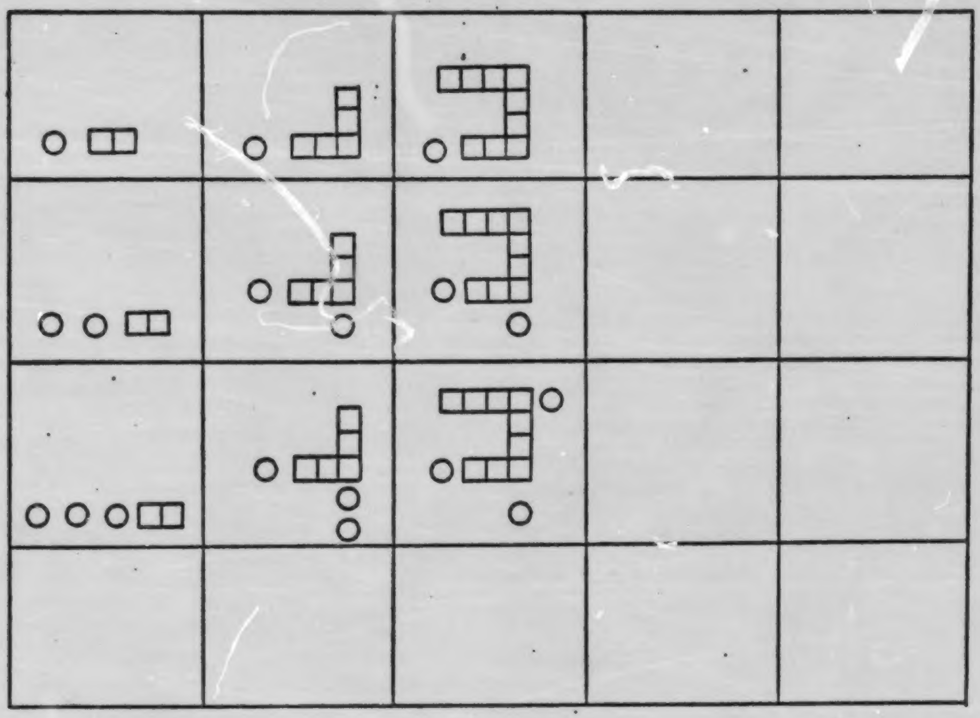
□	◍	△	▨	○	◬

In the first row there are as many differences as is possible between the objects. The first and second rows have to be filled in such a way that the neighbouring objects will be as different as possible one from the other. One solution consists in making three differences in the rows and two in the columns. It is impossible to make three differences everywhere. The diagram may also be filled by deciding that:

"neighbouring objects" will mean "neighbouring in a column". Is it possible to place the objects so as always to have three differences between two neighbouring objects? The problem is left to the reader.

To identify structures which underly certain patterns, it is best to use spatial arrangements since we can use arrow diagrams to emphasize the relevant relations.

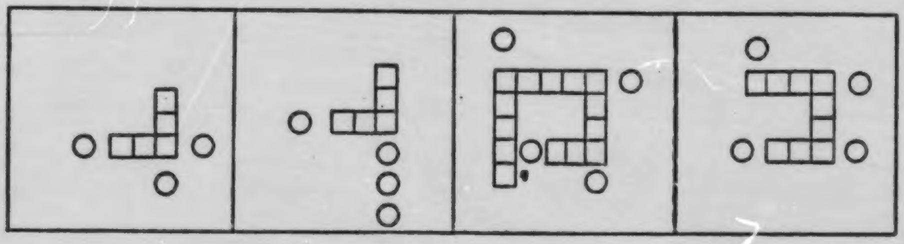
Here is a diagram of figures:



Here are other figures:

i) Identify the figures which have no place in the diagram.

ii) Put the other figures in the appropriate places.



iii) Fill in the empty spaces.

It is easy to see that the arrangement of the figures in rows and columns follows certain very specific rules. The squares cross each other from left to right, whereas the circles become more numerous as we proceed downwards. The squares produce spirals whereas the circles are lined up with the square elements introduced. If there are more circles than spiral elements, the circles are placed before the last element added.

The details of the exercise are not important. What is fundamental is the appreciation of the dynamics of a structure whose direction can be guessed from a small number of initial exemplars. The one who finds the rules inherent in such an arrangement, will have practiced the art of extrapolation. We shall see that this type of activity is at the very heart of the process of the learning of mathematics. In the preceding example we see a superimposition of two order relations. In the example showing differences, double entry tables have also been used, where relations of superimposed differences had to be considered.

A basic tool in learning mathematical ideas is the process of reasoning. The study of conditional situations forms the first steps of this process. To accustom the child to reasoning, embodiments can be given him in which he will discover the important properties of conditionals (or implications). After having understood what is meant by deduction, he will be in a better position to reason when dealing with more complex structures.

Here is for example a conditional situation:

In the figure, all the boys are wearing a hat. This means that the condition

"to be a boy"
implies
"to wear a hat"

We can see that it is not true that the condition "to wear a hat" implies "to be a boy", because there is a girl wearing a hat. However, the condition

"not to wear a hat"
implies
"to be a girl".

We can see that starting from concrete examples of a relatively simple kind, it is possible to learn to grapple with logical problems dealing with implications or conditions. For more details see the author's "Les ensembles et leur logique".

Pre-mathematics plays a very important part in the eventual full understanding of mathematics proper. It is quite possible that so few people understand mathematics because in school not enough attention has been paid to pre-mathematics. It is extremely probable that only by providing a great variety of mathematical materials that the understanding of mathematics by the general public can be made a reality. In the next chapter an outline will be sketched of how to bridge the gap between pre-mathematics and mathematics proper.

Chapter 3

The General and the Abstract

3.1 The particular and the general in everyday life

Nowadays we often hear of mathematics and its applications. Mathematics is often said to be very general; it is also said to be abstract. Very often, we are not too sure what is meant by such observations. In what way is mathematics general? What is meant by "general"? Is there a difference between the words general and abstract? Between the words abstract and symbolic? There seems to be quite an amount of confusion as regards the meaning of such terms.

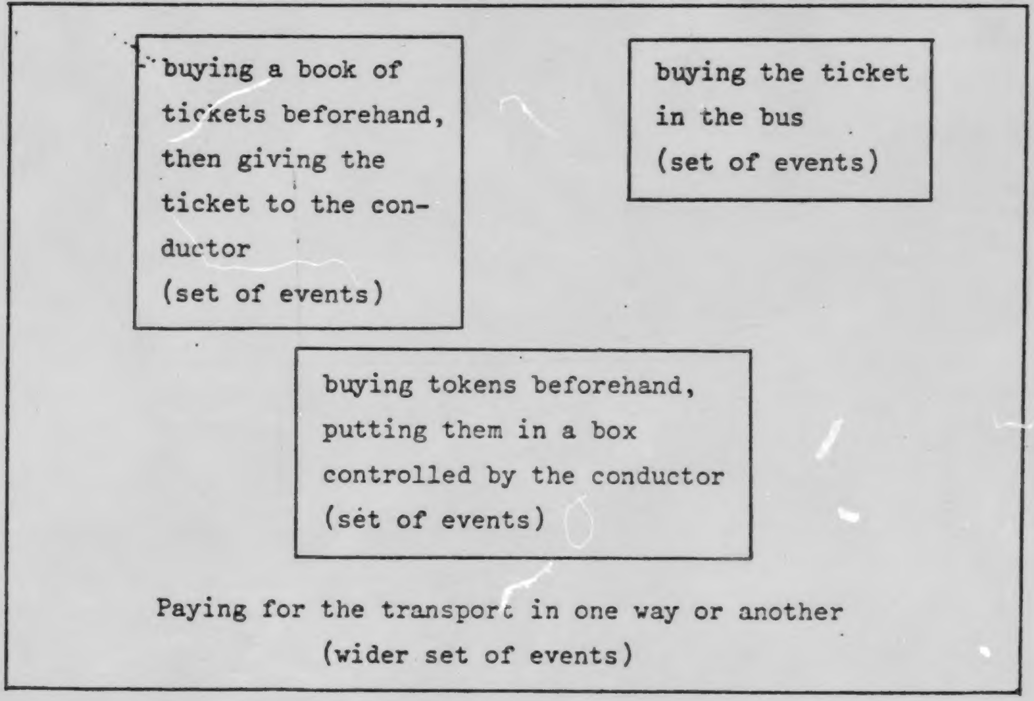
Take certain examples from everyday life. For example when two persons meet they say hello, they probably shake hands, or they may embrace each other, they may do other things, depending on their culture. In England, they say "How do you do"! In the States, they say "Hi"! In Australia, men shake hands, but never women. In Germany, they kiss women's hands but never the men's. In Italy, the men embrace each other. Imagine a person raised in one of these cultures and coming in contact with another culture whose customs and habits are very different. When two of these persons meet, will they be shocked by the completely different behavior encountered? Why would they be shocked? For Englishmen the "How do you do"? is related to an encounter between two persons. They cannot imagine things happening in a different way. They have not thought out, in a general manner, the whys and wherefores of suitable behaviour whenever one person meets another. They have only thought about it in a manner which is particular to their culture. They could not do otherwise, because they have never had other meetings than in their own culture. What would a general formulation of suitable behaviour be? What is always appropriate in any encounter? We should acknowledge the existence, indeed the importance of the other person, in relation to our own existence and importance. If the requirements for an encounter are formulated in this way, it can be readily admitted that there can be very different formulas or ways of realizing a communication of respect on the occasion of a meeting. Thus we shall not be shocked if in another culture we have to announce our arrival in a way very different from the one to which we have grown accustomed, the reason for this being that we have become aware of the more general idea of the requirements for an encounter. This generality has taken shape in our mind. We know that we need to communicate respect and politeness due to a person we have just met, and we are not necessarily any longer restricted to a particular rule or formula. Thus the particular ways in which we must behave in this particular culture are included in a whole set of situations in which a more general aspect of each situation is conserved and the particularities may be varied.

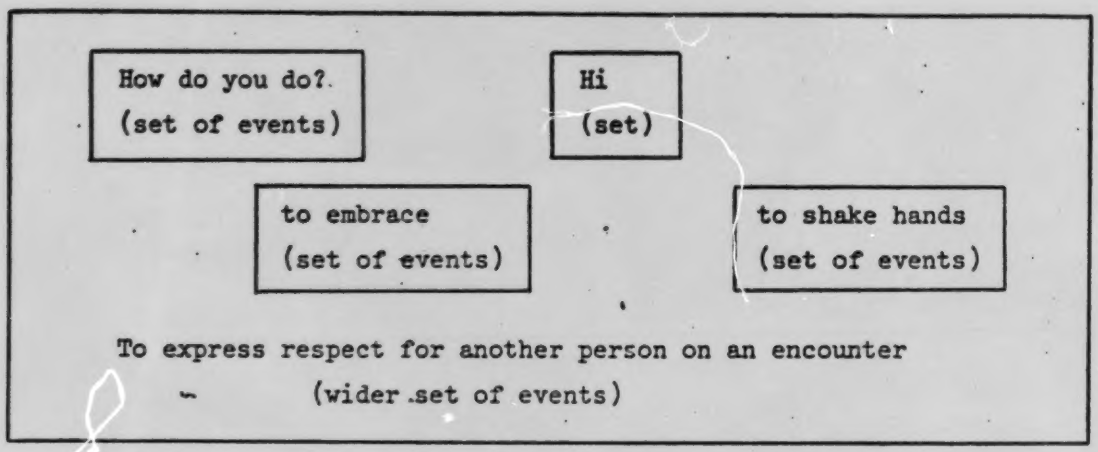
Take another example: when we make use of public transport we must usually pay for our ride. The methods used for paying can vary from one place to another. For example it may be necessary to buy tickets before getting onto the bus, or it may be necessary to have some particular form of change at hand because the conductor does not give out any change. We may have to get in by one door and get out by the other. If we are used to only one such way, we will certainly make a mistake while we get used to other means of paying for public transport. However, once we are used to many different ways, before using public transport, we will inform ourselves at the hotel about how everything is done in this city or town. Instead of thinking about the succession of "events", of getting in, looking for money, buying a ticket from the conductor, then giving it up before getting out, we simply say to ourselves that something is done to pay for the ride.

"Doing something to pay for the ride"
is evidently more general than
"getting in with money, buying a ticket, taking a ride and then giving up the ticket".

What is common to the examples quoted is that the passage from particular to general implies an extension of the sphere in which we are operating and in general implies a reformulation of certain operations in this more extensive sphere. This means that after having formed a class of "events", this class must be undone and replaced by another more extensive one, which contains at least a part of the operations made in the first sphere and which must now be made in a new and wider sphere. Certain properties of our operations will be lost but the important and essential aspects will be conserved, as for example, "paying for the transport". Indeed, even with this we could be very surprised. For example at "Man and his World" in Montreal it was not necessary to buy a ticket for the Island Express. This means of transport was free. Many must have wasted time looking for a place to buy tickets, if they did not know that public transport in this particular case was free. This means that this "event" does not enter into the general sphere "paying for public transport"; that is, it is not always true that one must pay for public transport, but the generalization consists in passing from a particular case of "paying for public transport" to the general case of "paying for public transport". Thus we go from one class of events to a wider class of events and in this wider class, the essential properties of the first class have been preserved. Evidently what we call "essential" remains to a certain extent arbitrary.

Here is a schematic way of representing the passage from the particular to the general.


buying a book of tickets beforehand, then giving the ticket to the conductor (set of events)
buying the ticket in the bus (set of events)
buying tokens beforehand, putting them in a box controlled by the conductor (set of events)
Paying for the transport in one way or another (wider set of events)



How do you do?
(set of events)

Hi
(set)

to embrace
(set of events)

to shake hands
(set of events)

To express respect for another person on an encounter
(wider set of events)

3.2 The concrete and the abstract in everyday life.

Let us now examine what might be meant by abstract. To illustrate generalization from a particular to a general case, we have already provided some examples. Let us now briefly examine the problem of the nature of abstraction. We come across nothing but abstractions in mathematics, but we meet abstractions also in many other spheres. Take for example the colors: red, green, black, etc. Red does not exist in the concrete world of objects. Only objects of a red color exist. Some are partially red, others are completely red. "Being red" is a property which can be possessed by an object. Red is not an object; "red" does not exist as an object. The property of being red serves to distinguish certain objects from other objects. These objects are red and the others are not. Any property which serves to distinguish certain objects from others, is an abstraction. A property is always an abstraction. An abstraction of an even higher order, is the notion of color. For example an object is red, but red is a color; therefore "to be red" is the property of certain objects and "to be a color" is the property of such a property, consequently, color or shape or size, etc... are abstractions of a higher order than red, triangular, small, etc...

How do we arrive at such an abstraction? For example take the concept of "red". If a set of objects is divided into two subsets one of which is complementary to the other in such a manner that all the red objects are put in the first and all those which are not red in the second, only the common property of the objects placed in the first set is considered. We realize that there is something in common to these objects and for this reason, we are in a position to put them together and to realize that they belong together. The fact that they belong together is made possible by their possession of a common property, that is, that of being red. If we were color-blind we could naturally not arrive at such a distinction and such an abstraction would forever be impossible for us to reach. Thus very probably certain abstractions are not accessible to certain persons because of certain deficiencies in their central nervous system. Fundamental mathematical abstractions are accessible to everyone except in pathological cases where there has been damage to the brain or in the case of certain psychological blockages.

In short, to abstract means to consider a certain property common to many objects or events and at the same time to reject the properties not relevant to the classification with which we are concerned. By means of this conservation and rejection, a new class of elements is formed. This discrimination is achieved through our intuitive realization of the communality of ideas, when we have constructed the class in question. There are two complementary aspects in the abstraction process: the positive aspect, which leads us to take into account certain common properties and the negative aspect, which leads us to ignore just what we do not wish to take into account, as regards the classification in question. For example, if a set of objects is divided into red ones and non-red ones, their shape, weight and size will not be considered as relevant. Only the color of the objects will be considered relevant in order to construct the set of red ones which we wish to distinguish from the set of non-red ones.

In general, we would not say that "red" is abstract, because the notion of red is so familiar to us. Higher order abstractions can be constructed from abstractions already constructed. For example, the abstractions red, yellow, green, black, white... all belong to a class we call colors. The abstractions triangular, circular, square... belong to the class of shapes. Thus notions such as color, shape, size, are abstractions of the second order.

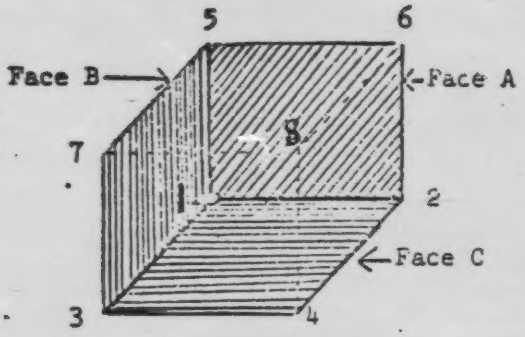
In mathematics we are continually passing to higher levels of abstraction, and this is one reason for the "difficulty" of mathematics. Certain abstractions are easier to construct than others. Certain similarities are quite striking, whereas others are more difficult to discern. Newton had to be a genius to discover that the fact that an object fell toward the earth was an "event" which ressembled the fact that the moon did not fall toward the earth, but moved along an orbit around the earth! These two events were not similar at first sight, and Newton had to eliminate a lot of "noise" before realizing that the two events were in fact similar from a structural point of view. This similarity was expressed in his theory on gravitation. Einstein discovered later that there was a similarity between the apparent irregularities in the orbit of the planet Mercury and the fact that light had the same speed relative to the earth, independently of its position in its orbit around the sun. These two pieces of information were still more difficult to put together in the same class because the two "events" were even more different from each other than Newtons apple and the moon.

Luckily, not everyone has to make such momentous discoveries, but everyone does make simpler discoveries because they are necessary. We have to set up categories; othervise every event would be considered different from every other event observed. If events are not classified, it will be impossible ever to make predictions after having observed a certain number of events. Evidently, every event in principle is different from every other event. But some of these events are similar enough for us to decide that after any event of that one class, certain consequences will follow. For example, we know that if we push a button to call the elevator, it will come. If it does not come, we say that it has stalled or that somebody is keeping it; we know it "should come". "To push on ~~a~~ button to call the elevator" is not just the one event but a class of events. An element of that class of events can be produced in the morning, at night or at anytime. The elevator can be called by us or by any other person. One of these events is equivalent to any of the others, and we know that the consequences of such an event is the arrival of the elevator. Abstraction in the form of classification is a common process, and one which is vitally important to the survival of an organism. The companions of our ancestors, thousands of years ago, who were unable to acquire enough skill in matters of abstraction and consequently of prediction, did not survive and never left any heirs! Thus, through natural selection, humanity has progressed to a situation where abstractions, necessary for survival, are more and more profound. To-day it appears very probable that only those who are able to use the tools of mathematics and of science to the best advantage will survive in the future.

3.3 The concrete and the abstract in mathematics.

At the level of relations and of more advanced mathematical notions, the principle of abstraction remains valid. To consider a mathematical structure as abstract, is to consider this same structure as being isolated from any irrelevant properties. For example, take the groups of blocks and the groups of sentences considered at the end of the preceding chapter. We constructed pairs of blocks and pairs of corresponding sentences. If we had abstracted the idea that in a pair there must not be one common property between the two elements of the pair, then the four pairs of the kind that can be built with the eight blocks, would immediately lead to the four pairs constructed out of our eight sentences. All there is to do is to change all relevant properties when passing from one element of the pair to the other. The essential part of this process is the same, whether we take the eight blocks or the eight sentences. The fact that in one case blocks of different size, shape and color have been taken and that in the other case sentences with different subjects, verbs and objects have been taken is no longer relevant.

Once the process of abstraction has been completed and another situation with the same structure is encountered, we should be able to consider four pairs constructed in the same manner as "old friends". We might take for example, the eight corners of a rectangular box.



Any given corner can be part of face A or not be part of it. A corner which is part of face A has the property A. Property B and property C will mean, being a part of face B or of face C, respectively.

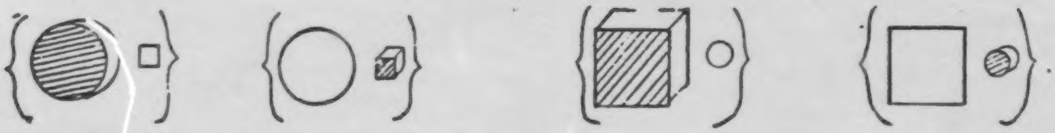
Corner 1 has the three properties A, B, C. We will have

1 = ABC, 2 = AC, 3 = BC, 4 = C, 5 = AB, 6 = A, 7 = B,

8 = neither A, nor B, nor C.

The pairs (1,8), (2,7), (3,6), (4,5) all contain elements different one from the other from the point of view of all three properties. It is interesting to note that the sum of the two numbers used in a pair always turns out to be 9!

The abstract structure which determines these pairs is the same as that which determines the pairs



where the elements of each of the pairs are different one from the other

i) in shape
ii) in thickness
iii) in size.

"To be on face A" corresponds to "being circular"
"To be on face B" corresponds to "being thick"
"To be on face C" corresponds to "being large".

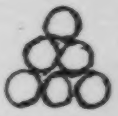
If abstraction has been made of properties relevant to the construction of the pairs, we should see that our four pairs are constructed in the same way from a structural point of view, whether we have four pairs of blocks having no common property or of four pairs of sentences whose elements have no subjects, verbs or complements in common. We can ignore the fact that in one case we speak of shape, thickness and size, and in the other of corners situated on one face or another, and that in the case of sentences, we speak of the identity or the difference of the subjects verbs or complements of the sentences. We keep the common kernel of the activity which consists in building pairs whose elements have neither one nor the other, nor the third property in common. This is the abstract kernel of the three activities. The reader might try and find a fourth example having the same structure as that of the examples already mentioned. If he is able to find one, he can be sure to have understood the abstraction in question.

We have seen that when we generalize, we pass from a set of objects and events to another set of objects or events which is more extensive than the first set, while conserving certain essential properties which interest us, as we pass from the first set to the more extensive set. What happens when we abstract? There are no sets of objects at the start, there are only certain elements. In the case examined, there is an element of the eventual set of "events", that is, the four pairs of blocks. This is the first element which will have to belong to a set yet to be constructed. The second element is the four pairs of sentences; another element is the four pairs of factors of 30, still another is the four pairs of opposite points, that is, the extremities of the principal diagonals of a box. If we can continue in this way and form other elements and as many as we wish, this means we have realized what is the common property of these elements. We have constructed the essential and fundamental property by virtue of which the elements were judged similar. Indeed, a class has been constructed from these elements. Abstraction is a creation. The class has to be created from the elements by realizing common, relevant properties possessed by these elements. On the other hand, generalization is a process in which we go from a set already constructed to a larger set to be constructed, while keeping certain properties which link the elements of the same class to one another.

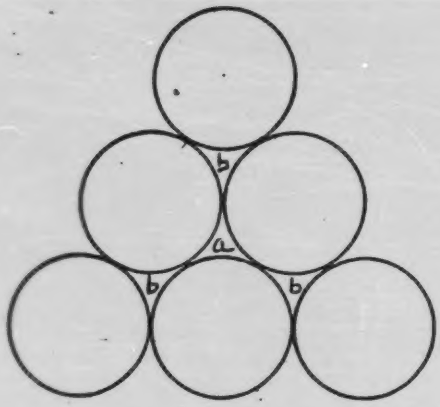
3.4 The Particular and the General in Mathematics

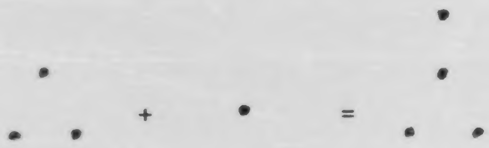
Let us now resume our discussion on the general and the particular. We have seen that passing from the particular to the general is the process of generalization; the opposite process, that is, going from the general to the particular could be called the process of particularisation; in mathematics it is often called substitution, because something which is general is replaced by something more particular. For example, a variable in a formula can be replaced by a value of the variable; this means that we go from the general formula to the particular application of the formula. However, we can sometimes arrive at a general formulation from certain particular cases, and this is what is called generalization, going from the particular to the general.

Let us suppose that children have studied triangular numbers and square numbers. A square number is obtained by multiplying a natural number by itself; a triangular number is obtained by adding the first natural numbers starting with one and stopping at a certain number, For example, one plus two plus three plus four. Consequently, ten is a triangular number. However, ten is not a square number, because it can never be obtained by multiplying a natural number by another natural number. However, sixteen is a square number because four times four gives sixteen. By working with triangular numbers and square numbers, children realize all of a sudden that if they add two successive triangular numbers they always obtain a square number. In cases where it can be checked, we find that this is what happens. One may wonder if this is what always happens. Take, for example, the number three, which is the second triangular number, and the number six which is the third. Their sum is nine, which is a square number. The third triangular number is six. If it is added to the fourth triangular number, which is ten, we obtain sixteen. We have just seen that sixteen is a square number, Is this the case only for small numbers purely by"accident"or is it always so for all numbers? Such ways of passing from particular cases to a general law would be a generalization. We could generalize by induction simply by positing a hypothesis. But we still have to be convinced that our hypothesis is valid. We could ask whether it is possible to prove in one way or another that our supposition also holds good for big numbers. The hypothesis cannot be "checked"for all numbers, because there are too many. We must face the problem of proof if we wish to know whether what we suppose to be true in general is indeed true in general. Evidently, it remains to decide in principle what kind of procedure will be accepted as a proof.

Let us try the following procedure: take some beer-mats that you get in bar saloons. Three can be put together so that the circles touch; we will have a triangle: 1 plus 2. Then we could still put three alongside, like this:  which would make: 1 plus 2 plus 3. By making triangles with mats, making sure that the circles are as close as possible to each other, we will always get triangular numbers.

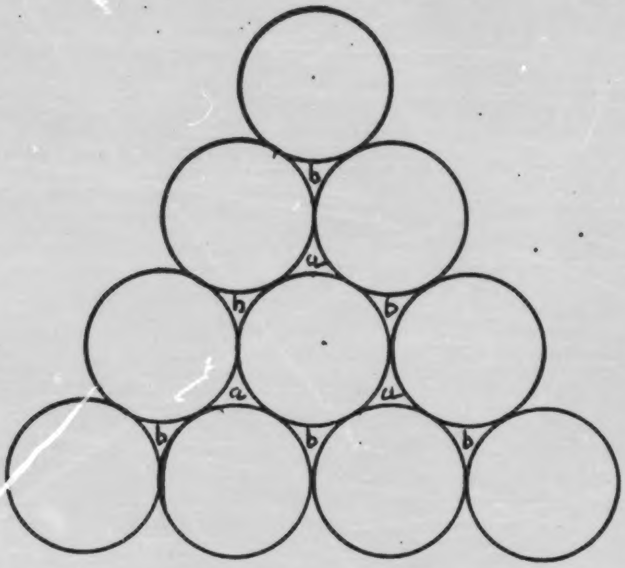
Now let us cover the "holes" which have been left between the mats with other but smaller mats. For example, if six mats have been used to make a triangle, there will be exactly four holes which can be covered. Moreover, there will be two types of hole. One type will be found immediately below a circle, the other immediately above a circle. There will be three holes of the below or 'b' type, and one hole of the above or 'a' type, as shown in the diagram.





3 + 1 = 4

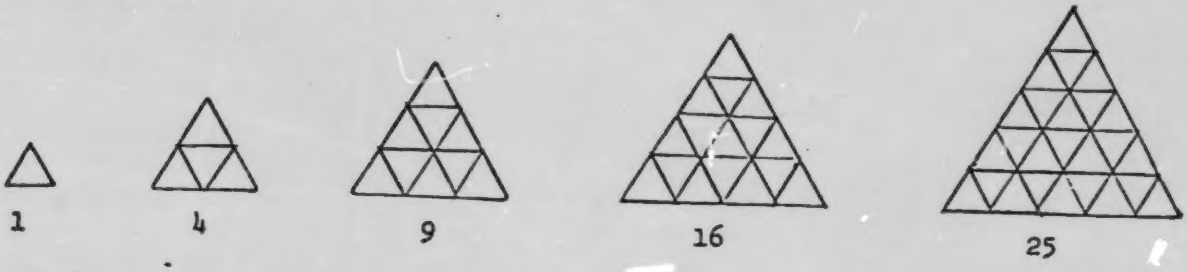
Adding the next row of circles will give the diagram



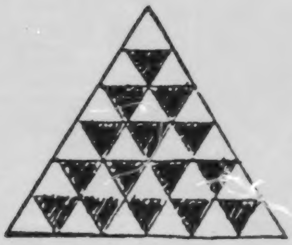
Here we have six holes of the 'b' type and three holes of the 'a' type. Adding four circles has added five holes, three 'b' type holes and two 'a' type holes.

What happens if we add another row? We shall then have added five circles and get four more holes of the 'b' type and three more holes of the 'a' type, seven extra holes in all. Next we should get nine extra holes, five of the 'b' type and four of the 'a' type. Consequently, as we proceed with our construction we always add one number for the 'a' type holes and one more than this number for the 'b' type holes. Given two "neighbouring" numbers, one is always odd, the other even, and the sum of an even number and an odd number is always odd. We are therefore always adding an odd number. Also, every time we proceed to the next stage of our construction, there are two more holes to fill than there were before, with the result that we add an odd number which is two more than the odd number of the previous stage.

Consequently we always add consecutive odd numbers. We can see that in doing this, we are always bound to make squares; we only have to put triangles together to see that a square number of triangles is always necessary to make a larger triangle. After the first triangle, three have to be added to make a larger triangle consisting of four small triangles; then five must be added to make the next triangle, which will contain nine small triangles, and so on.



We will also see that the triangles which appear "the right way up" will replace the mats used at the beginning of our mat-construction and the triangles which are upside down will replace the mats used for covering up the "holes" left between the mats.



21 + 15 = 36 .

21 = 1 + 2 + 3 + 4 + 5
15 = 1 + 2 + 3 + 4

Thus, we believe to have "proved" that if we add two successive triangular numbers we shall always obtain a square number.

3.5 Towards the more General

Up to now we have seen that the set of initial events was contained in the set of events which were the object of our generalization. There are other kinds of generalization where the initial set is not, strictly speaking, included in the generalized set but the generalized or extended set possesses a structure very similar to that of the primitive set. Take, for example, a cycle containing three elements and compare it with a cycle containing five elements. Here is a game we can play with three different positions of the arm.

The arm up, the arm in a horizontal position and the arm down will be states of the game. One person will represent the state of the game and another will represent the operator. There will be three operators: the horizontal operator, the "up" operator and the "down" operator. When the operator puts his arm in the horizontal position, the "state person" remains in his position, he does not change; when the "operator person" lifts his arm, the "state person" makes the following movements:

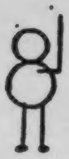
if his arm is horizontal, he lifts it
if his arm is up, he brings it down and
if his arm is down, he brings it to the horizontal position.

If the operator brings his arm down, the "state person" does the opposite:

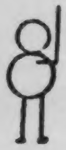
from the horizontal position, he brings it down,
from the "down" position, he lifts it up and
from the "up" position he brings it to the horizontal position.

Thus we have made up a game of three elements.

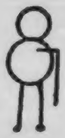
Here is a possible realization:



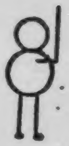
State



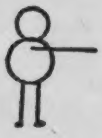
Operator



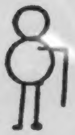
State



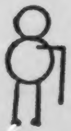
Operator



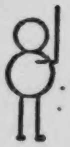
State

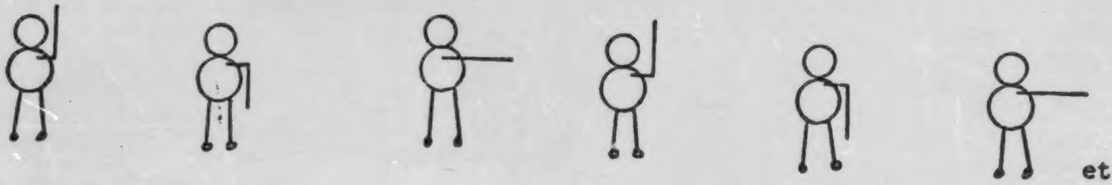


Operator

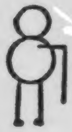


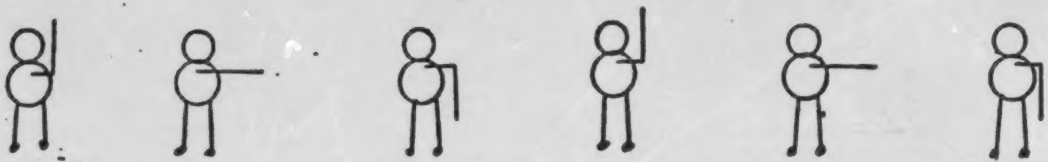
State

If the same operator  is applied all the time, we shall have, for example, the following states:



etc...

On the other hand, by applying the operator  we shall get these states:



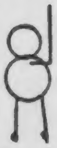
etc...

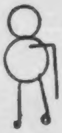
We use the same kind of operators when we wait for a bus. Suppose that bus A goes by the street corner at 5, 25 and 45 minutes after the hour, and that bus B passes every 40 minutes, for example at 8.05, 8.45, 9.25, 10.05, 10.45 and so on. If we have just missed a bus and must wait for the next one, in the case of bus A we have the following possibilities:

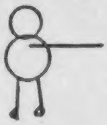
missed bus	next bus
05m after the hour	25 m. after the hour
25 m. after the hour	45 m. after the hour
45 m. after the hour	05m. after the hour

On the other hand if we have just missed bus B, we will find ourselves in one or the other of the following situations

missed bus	next bus
05m after the hour	45 m. after the hour
25 m. after the hour	05m. after the hour
45 m. after the hour	25 m. after the hour

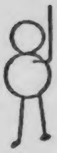
We would see that the first situation corresponds to the operator 

the second to the operator 

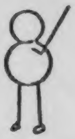
When we catch the bus, we are in the situation of not having to wa[illegible] that is in the situation represented by the operator 

We can compare this game to a game of five elements with the following positions of the arm. Let us call them:

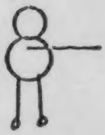
two up
one up
neutral
one down
two down



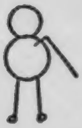
2 up



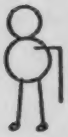
1 up



neutral



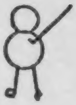
1 down

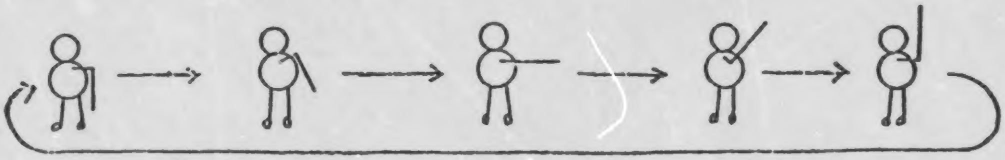


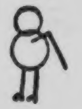
2 down

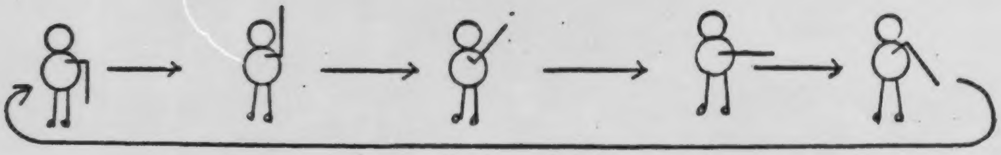
In this game, there will also be a person acting as a state and another as an operator.

"1 up" will be the operator of the movement of the arm from the position "2 down" to the position "1 down", then to the neutral position, then to the position "1 up" then to the position "2 up" and then to the position "2 down", The opposite of this order will be "1 down", which will give a cycle in the opposition direction.

The "1 up" operator or  is a movement along the following cycle:



The "1 down" operator or  will result in a movement along the following cycle:



The order or command "2 up" will be the movement from the "2 down position" to neutral position. From this to "2 up" position, from this to "1 down" position, from this to "1 up" position, from this we return to "2 down" position.

The opposite of this cycle will be engendered by the command "2 down".

We can now see that the two games are indeed similar. In both cases, there is a neutral, in both there are cycles in one direction and in another; in the game of three there is only one cycle in one direction and one cycle in another, whereas in the game of five, there are two cycles in one direction and two cycles in another. In this sense the second game is an extension of the first but from a structural point of view we can see that the first game is really not contained in the second one because there are no elements in the three cycle of which are also elements in the five-cycle, because all the cycles are five-cycles in the five-game whereas all the cycles in the three-game are three-cycles. The three is replaced by five when we pass from the three element game to the five element game. Now to generalize the game, we could think of a game with seven elements. Evidently, there will be three cycles in one direction and three cycles in another and there will be a neutral element, etc. Evidently also in a game with an even number of elements there will be no symmetry with regard to the neutral element as has been observed in the case of an odd number of elements.

Why is it necessary for the pedagogue and also for the parent to have a certain knowledge concerning the abstraction process and the generalization process? Today it is becoming more and more necessary for the population in general to be familiar with mathematics. The increasing complexity of our society, in technological, economic and social organization, requires much more specialized man-power. This specialization of man-power is accompanied by a literature, a culture and even jokes and stories which are part of this new culture. A citizen now finds himself in new circumstances, which are totally different from those existing 15 years ago; the teacher and the child must understand certain learning processes so as to be able to learn more easily what is necessary to adapt to this new world. Both teachers and parents should be able to help children with their learning as they grapple with mathematical problems at school and at home.

3.6 Simple to Complex or Complex to Simple?

Since there are certain mathematical notions and techniques that should be learned, we can for example present the problem: "What is the optimum order in which to learn these notions and techniques?" We have supposed until recently that it is always more efficient to learn a simple structure followed by a more complex structure. There is today some evidence that the contrary may be the case in some instances. This means that it may be more efficient to start learning a more complex structure before a simpler one, if it is necessary to learn both.

The reader might try experimenting with the game we have just described. He could take three or four children and teach them the game with three elements, and the following day, the game with five elements. He could then take another group of children and try to teach them the game with five elements first and then the game with three elements. Evidently this will not be a controlled experiment of a scientific nature but it is quite possible that he will find that those learning the game with five elements before the game with three, will learn with less difficulty than those who learn the games in the reverse order. One might think that the more natural order would be going from simple to complex rather than from complex to simple. If we think it over, we might realize that in the simple case there is not much to learn, consequently we do not see what the structure consists of; whereas in a game with five elements the child will see more clearly the elements of structure to be learned and thus he will learn with more facility the structure with three elements after having learned the structure with five. The child who has learned the structure with three elements will learn the structure with five with more difficulty because he has limited himself to a simple case and will have some difficulty in getting out of the simple game. If he started with the more complex case he would not face this difficulty. It is often said in psychology that in certain cases there is a positive transfer from complex learning to simple learning and that sometimes there is consequently a negative transfer from the learning of a simple structure to the learning of a more complex structure.

There are other spheres where priority of the complex over the simple is acknowledged by the experts. For example, the dynamic reading method supposes that the unity of an utterance is primordial, and that we must start with "sentences" and analyse them afterwards. The complexity of the sentence situates the learning matter in a context which becomes easy for the child to understand. He learns how to subdivide a sentence, he reaches the "words", the "syllables" and finally to the letters and their links with the sounds he hears when he speaks.

The method of total immersion in the learning of a second language is based on the same fundamental hypotheses. Complex activities of a mathematical nature provide a kind of immersion to the child and allow him to place these experiences in a context, rather than learn isolated "facts".

One must not imagine that the order of complex to simple is always more efficient than that of simple to complex. The choice evidently depends on the degree of complexity as much as on the difference between the two complexities. In the psychology laboratory, in cooperation with the International Study Group for Mathematics Learning, experiments are being conducted to determine with more precision what degrees of complexity are optimally adapted to begin learning and in what order the complexity should vary afterwards. Such an experiment can last from six months to a year, consequently we must wait a few more years before it is possible to give more precise information in this area. In general Research Councils and Foundations have not yet understood the importance of this type of research. Since the juries who examine the requests are often not well informed, it is difficult to begin any large-scale research of the kind. In spite of these difficulties there are already about twenty experiments being done in about twelve countries where local universities or private agencies have contributed.

We hope that after the publication of the first results it will be easier to obtain the support of government agencies for such research which is of prime importance for the reform of modern mathematics teaching.

Chapter 4

The Passage from Pre-Math to Mathematics

4.1 The first Steps

We saw in the description of pre-mathematics that this activity on the part of the student consisted of certain structured games following free play, which was itself an initial adaptation to the child's environment. The first step in learning is always a first adaptation. It is what has been called free play. The second step consists in determining the constraints of the situation on the one hand and imposing constraints at will on the other hand. Manipulation of imposed constraints is what is called a game. If the imposed constraints follow certain mathematical structures, then the games will introduce the child to situations which are similar, from a structural point of view, to the mathematical properties they will have to know later.

We saw at the end of chapter 2 that it was possible for children to play alongside one another certain games which had a common structure. The process of abstraction consists precisely in realizing what the common properties are which in effect make these games similar. A very important step which follows the second step would consequently be the comparison of these structural games. Some of these games have the same structure whereas others will be found to be in relationships of inclusion or of exclusion, to each other. The realization of such comparisons will be part of the third stage in learning.

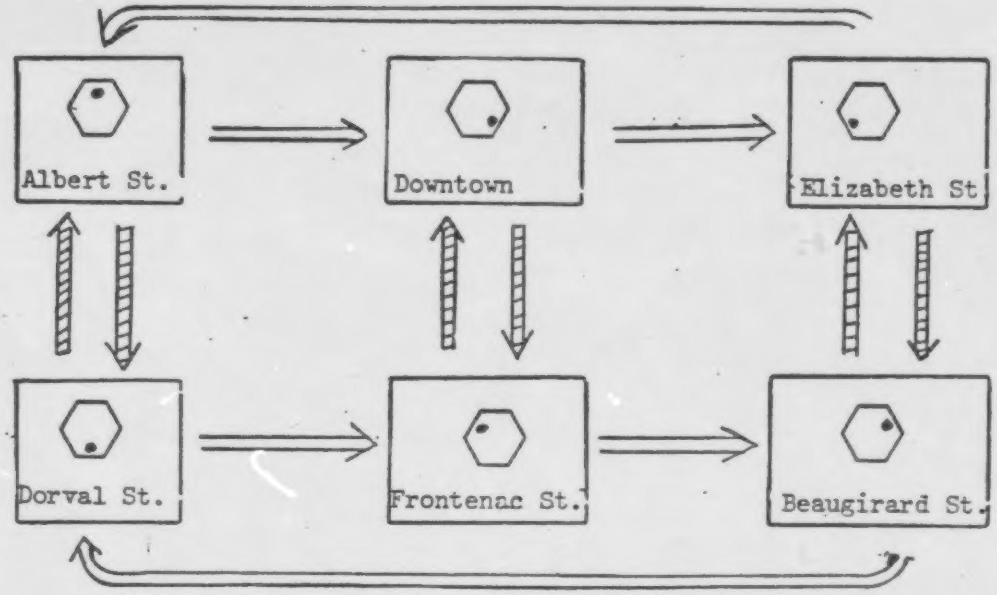
After having extracted the common abstract contents of a number of different activities, it is time to use some representations. This is where we really pass from pre-mathematics to mathematics. When we play games with objects, we are not doing mathematics; we begin when we start comparing a structured game with another. It is the realization of such a comparison in detail which constitutes the first steps toward the abstraction of mathematics. Thus once the child has had a hint of the nature of a situation, it is time to give him some coding systems. A first coding system should undoubtedly be visual, that is, spatial. We have already seen examples of such spatial systems; they were arrow diagrams. An arrow diagram can incorporate the abstract structure it represents. The child is able to see all at once all the details he has discovered during the games and during his efforts in comparing these games. He will have classified certain games as being of one kind and will say to himself: "Here is a way of representing all the games with rules of the same sort". It will usually be an arrow diagram of all the games of that kind.

4.2 The analysis of the six-cycle

Here is an example: Take a regular hexagon cut out of cardboard and pin it to any surface with a thumb-tack at the centre of it. Draw the outline of the regular hexagon on this surface, then make either a half turn, or a third of a turn in one direction, for example, clockwise. Thus by doing a certain number of turns either of the first or of the second kind, the hexagon can always be brought to any position in which it occupies the space inside the outline drawn. There are obviously six possible positions subject to this constraint. We can see for example that, to make the hexagon turn 60° clockwise, we must do a third of a turn twice and half a turn once. We thus find certain similarities between different series of movements of half-turns and third-of-a-turns. For example, a half-turn, a third-of-a-turn, followed by a half-turn and a third-of-a-turn could be replaced by two one-third-of-a-turns, one after the other; or else one third-of-a-turn followed by a half-turn followed by a third-of-a-turn followed by a half-turn, followed by a third-of-a-turn, always brings us back to our starting point, etc ...

Then, we can play another game which has the same structure. For example, we could think of a town with a circular bus, route no. 1, which leaves downtown, stops at Dorval Street, then at Elizabeth Street, then at Frontenac, Albert, Beaugirard, and returns downtown. Suppose there is bus no. 3 which is an express and stops only twice during its whole run. This means that after any particular stop, it does not stop at the next one nor at the one after that, but it stops at the third stop. Number 2 will stop at every other stop. For example, if we take no. 3 downtown, it will stop at Elizabeth Street, then at Albert Street, then downtown. If on the other hand we take it at Dorval Street, then it will not stop at Elizabeth Street but only at Frontenac Street; it will not stop at Albert Street but only at Beaugirard Street and then it will go past downtown without stopping and stop again at Dorval Street. Thus, if we want to go from downtown to Dorval Street by bus we will have to take bus no. 2 and stop at Albert Street, where we will take bus no. 3 which does not stop at Beaugirard Street, nor downtown, but at Dorval Street. Thus, with a transfer we can come from downtown to Dorval Street. We can easily see that the role played by bus no. 3 is the role played by the half-turn in the hexagon game and that played by bus no. 2 corresponds on the other hand, to the role of the one-third clockwise rotation.

To represent these two games or any other game of this nature, we can show the six states in the following manner:



The white arrow represents the cycle completed after three applications of the operator which produces it. The striped arrow represents the two-cycle, completed after two operations. This last one corresponds to bus number 3 or to the half-turn in the case of the regular hexagon. The first one corresponds to bus number 2, or to the one-third-of-a-turn in the case of the regular hexagon, Thus, this arrow representation allows the student to see all at once all the possible properties of the operators in a visual form, as far as equivalences are concerned. We shall see, for example, that to do one stop on **bus no. 1 comes to the same thing as taking no. 3, transferring to no. 2 and alighting at the second stop.**

4.3 The equilateral triangle

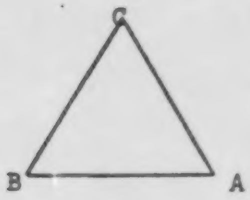
We shall now describe another game with an equilateral triangle. We can draw its outline after having placed it on any surface. We can see that if we place the triangle with one side up there are three possible positions in which it will occupy the space, namely the space inside the outline. If we turn it over to its reverse side, there will be three more possible positions. Thus, we can choose two different operators to produce these different positions from any initial position. The first operator could be for example:

one third of a turn, clockwise

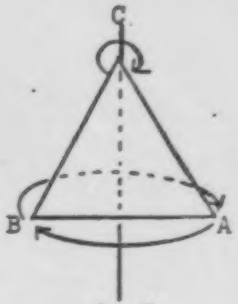
and the second:

turn the triangle back to front by holding it along the axis shown in the diagram.

The first operator which makes the third of a turn could be called X, the second one Y.



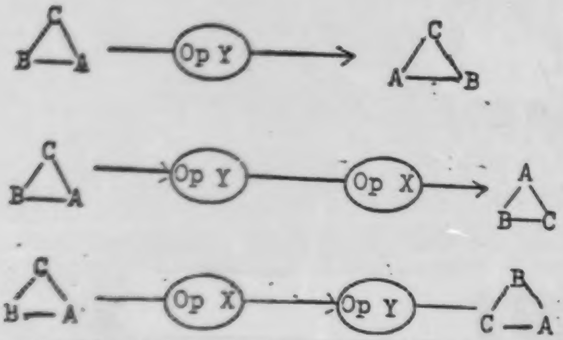
Op X



Op Y

We can see that it is always possible to move the triangle from any one position to any other position, in not more than two moves.

Here are the means by which this can be done:



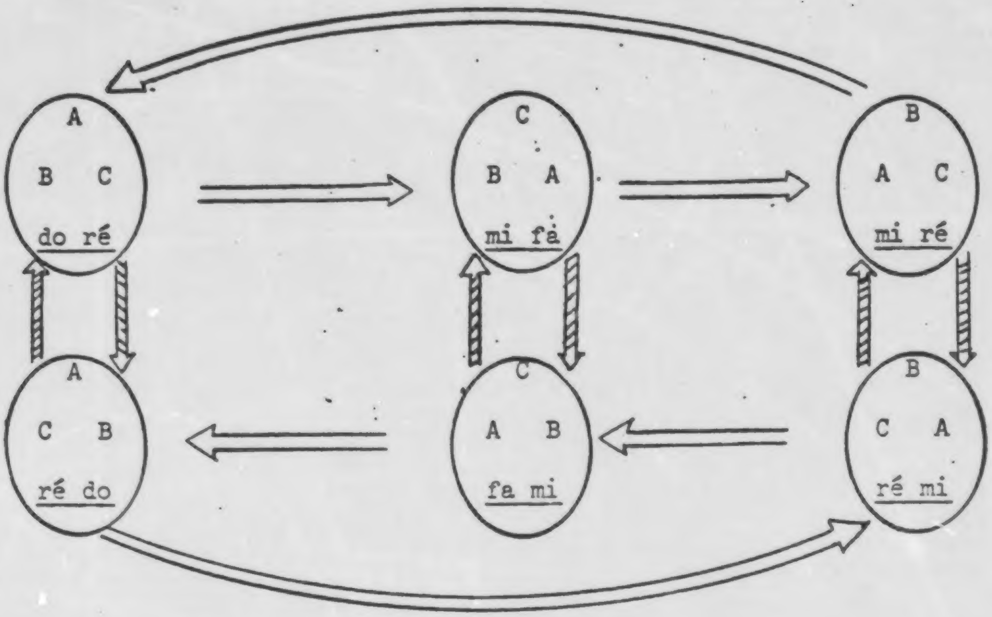
A similar game can be played with musical notes. Take for example do, ré, mi, fa. At the piano, exercises are often performed going up and down the scale do, ré, mi, fa. For example: do, ré, mi, fa, etc ... This means we are going up and down. Subdivide the exercice into pairs of notes. For example, if we begin with do, re, when we continue we come to mi, fa, then to mi, re, then to do, re. This means that to continue is an operator which brings us back to the initial state, that is, in this case, to do, re, after three applications of the operator. We can also play with another operator, for example:

"Reverse the order of the notes".

If we have come to do, re, to reverse means to pass to re, do. If we have just said mi, fa, to reverse means to pass to fa, mi, etc ... Let us see for example what happens if we mix the two kinds of operators. Take the chain of operators:

continue - reverse - continue - reverse

we finally come back to the initial state. For example start with do, re; we continue, we come to mi, fa, we reverse we come to fa, mi, we continue, we come to re, do, we reverse, we come to do, re, which was our initial state. It can be verified that if we do the operations X, Y, X, Y by starting at any initial state of our triangle we will come back to our starting point. This suggests that the two games have some common properties. Indeed, from a structural point of view, the two games are identical. They can be represented by means of a common arrow representation such as shown in the following diagram:



The diagram shows that the third of a turn or the operation "to continue" can be represented by the white arrow, whereas the "flipping over" operation in the triangle or that of reversing the musical notes, can be represented by the striped arrow. Thus we see that the arrow representation gives us in one picture a general view of all the possible properties of this game which have to do with a series of successive operators. We can also see that if we look at the two arrow representations, that is, that of the hexagon and of the bus routes on the one hand and that of the equilateral triangle and of the musical example on the other hand, that there is an important difference. In the first representation, the white arrows go round in the same sense whereas in the second, they go around in opposite senses in the two cycles. Thus we immediately see a fundamental difference between the two by looking at the appropriate representation and it will be more difficult for us to confuse them than if we just spoke of them vaguely.

The technique of comparing two games of identical structure consists in finding arrow diagrams to represent a whole class of these games, all of which possess the same structure. We shall eventually have to compare games essentially different from each other from a structural point of view.

The structural reason for which the games studied are not similar has been found; thus, exercises in studying similarities and differences introduced in a previous chapter can be developed. We are dealing here with quite complex structures. We are facing problems of discrimination at a conceptual rather than at a perceptual level.

4.4. An Order Relation

Let us now go on to another kind of representation. Take for example the comparison of a number of possibilities from the point of view of our personal preferences. Take the following example: I do not as yet know if I should or should not pay a fine of $100.00. My wife has invited my mother-in-law but I do not know if she has accepted the invitation. I may go and visit my friend some night next week, a friend I like very much. Thus, there are 8 possibilities.

We may use the symbol A for the paying of the fine and non-A for the non-paying of the fine.

Let us use B for the eventuality in which my mother-in-law has accepted the invitation, by non-B that she has not accepted it.

Finally, let us use F for the eventuality that I should be able to visit my friend and by non-F the possibility of my not being able to do so.

Thus, we have the following possibilities:

A B F, A B nonF, A nonB F, A nonB nonF, nonA B F, nonA B nonF, nonA nonB F, and nonA nonB nonF.

How could we arrange these eight possibilities in order according to our personal preferences? The order in which we place them depends on our financial conditions, on our affection or otherwise for our mother-in-law, and on our affection for the friend we intended to visit. Suppose we are not financially well off so that the $100.00 fine would represent a near catastrophe. Let us further suppose that we do not like our mother-in-law very much, and would prefer not to see her, and wish she would not accept the invitation; but her coming would not be as big a catastrophe as the payment of the $100.00 fine. Also, let us say that we would not be embarrassed saying "no" to our friend, but that not to have our mother-in-law over seems more important than to pay a visit to our friend next week. At this moment, what would be the most preferred eventuality? Evidently

non A, non B, F

would be the best. After that would come

non A, non B, non F

because this is slightly worse than non A, non B, F, because we nevertheless prefer to visit our friend to not visiting him. After this, would come the eventuality

non A, B, F, then non A, B, non F.

After these

A, non B, F

then

A, non B, non F

then, I would have

A, B, F

and the worst eventuality would be:

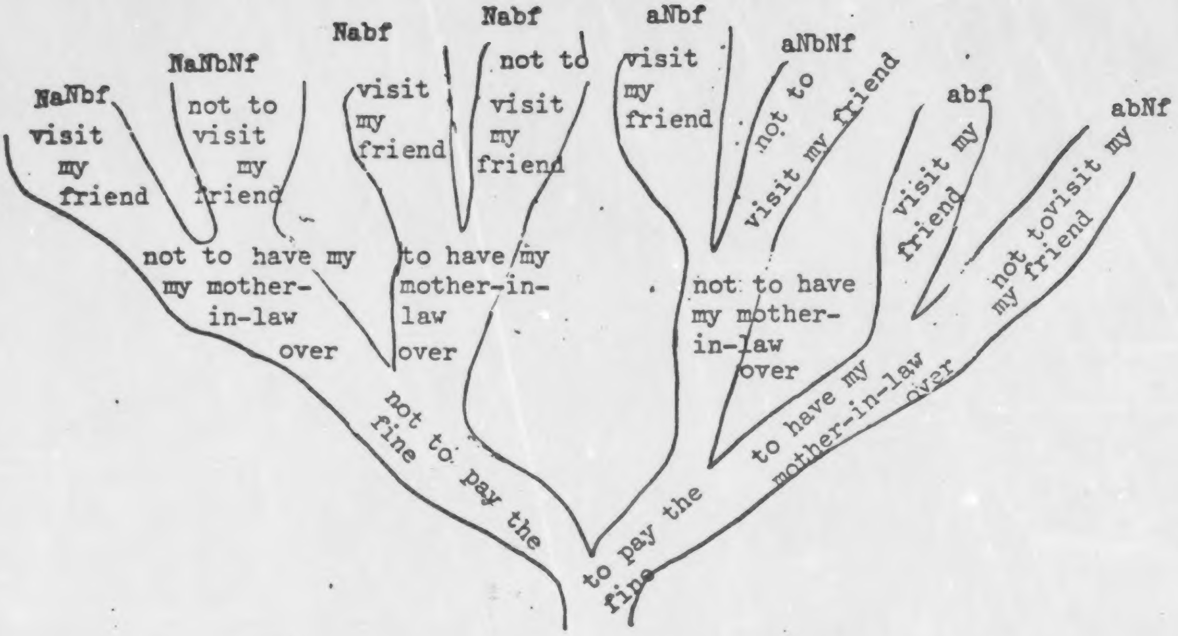
A, B, non F.

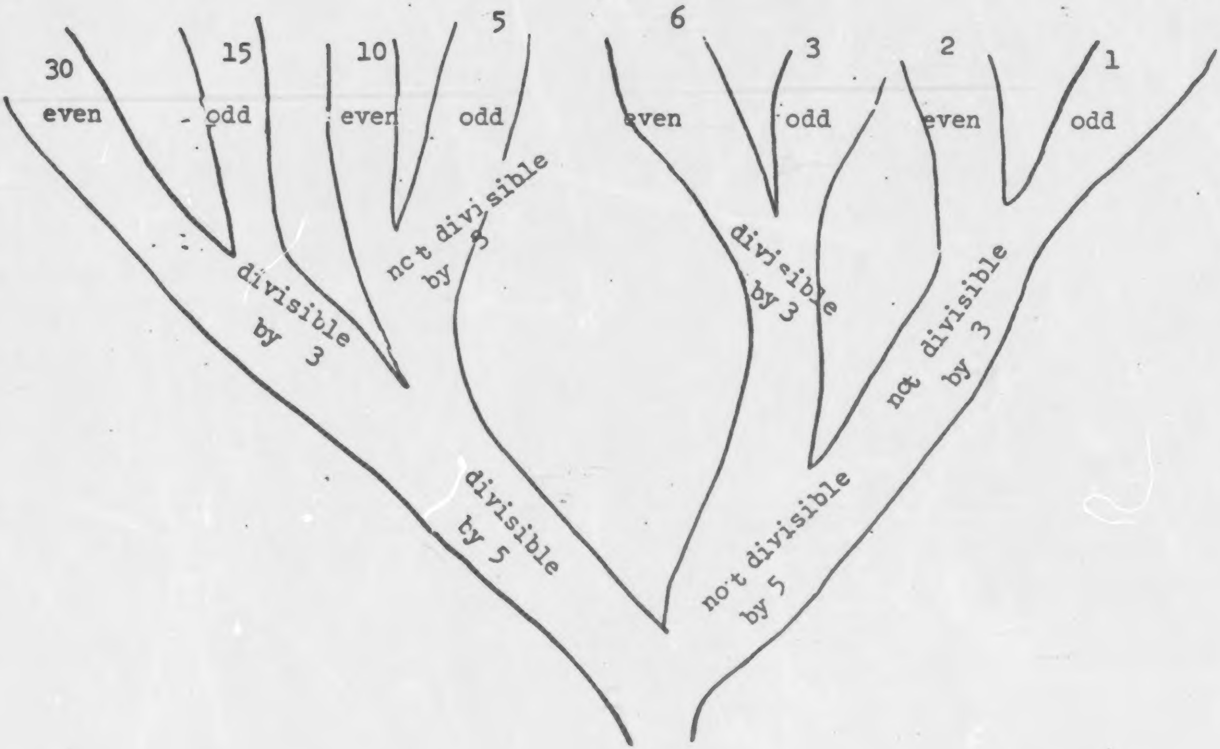
We could also make a similar comparison between the factors of 30; we could easily say that we liked the numbers with a factor of 5; then, if we have a choice, we take a number divisible by 5, but if the numbers available are not divisible by 5, then we should rather choose numbers divisible by 3 rather than any that are not. But divisibility by 5 is regarded as more important than divisibility by 3; if we cannot choose the numbers divisible by 5 or by 3, then our choice should be an even number rather than an odd number. Which would be our order of preferences? It will be the following:

30, 15, 10, 5, 6, 3, 2, 1

The first four numbers are divisible by 5 and the last four are not. Consequently, the first criterion will subdivide the eight numbers into two parts; the first part will contain those that are divisible by 5 and the second those which are not. Now we can choose between the numbers divisible by 5. But which one shall we actually choose? We can choose between 30, 15, 10 or 5. We like divisibility by 3. This means that we will choose either 30 or 15; or, if we have to choose between 6, 3, 2 and 1, then we would choose either 6 or 3 because we like divisibility by 3. But we must eventually make the final choice. If we had to choose between 30 and 15, we would choose the even number because our preference goes to the even numbers other things being equal; if we must choose between 10 and 5, the choice will clearly be 10; between 6 and 3 it will be 6; between 2 and 1 it will be 2. Thus, we have established an order of preferences by a set of rules similar to that of the preferences between the combined possibilities in which we must pay a $100.00 fine or have mother-in-law over for dinner or visit a friend.

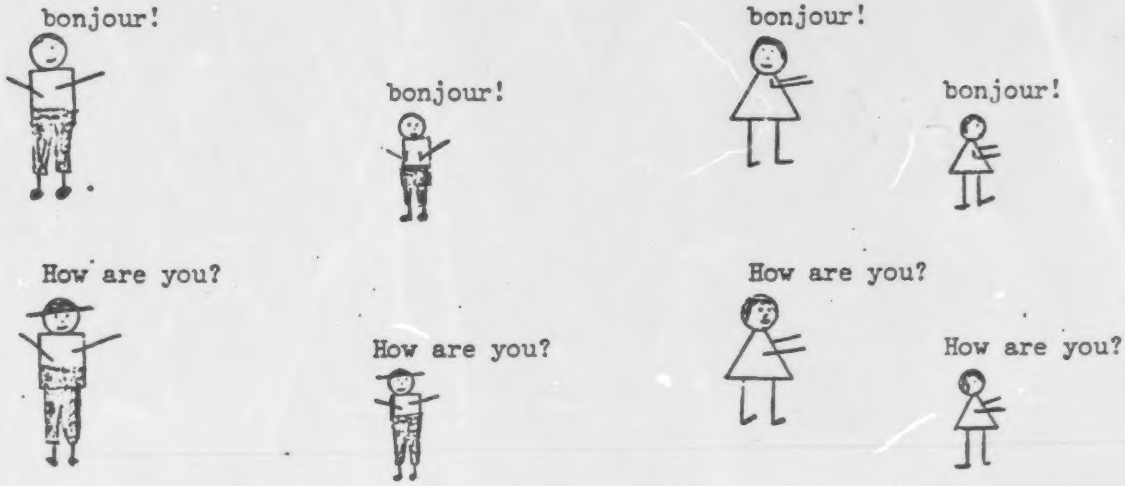
How could we show the common properties of these two types of evaluation of preferences? We could draw a tree-structure whose endpoints have empty spaces and in which we would write the different eventualities considered in the first case and the numbers considered in the second. Here are two tree-structures: the first one is filled with the eventualities of our first example and the second with our numbers,


NaNbf
visit my friend
NaNbNf
not to visit my friend
Nabf
visit my friend
Nabf
not to visit my friend
aNbf
visit my friend
aNbNf
not to visit my friend
abf
visit my friend
abNf
not tovisit my my friend
not to have my my mother-in-law over
to have my mother-in-law over
not to have my mother-in-law over
to have my mother-in-law over
not to pay the fine
to pay the fine

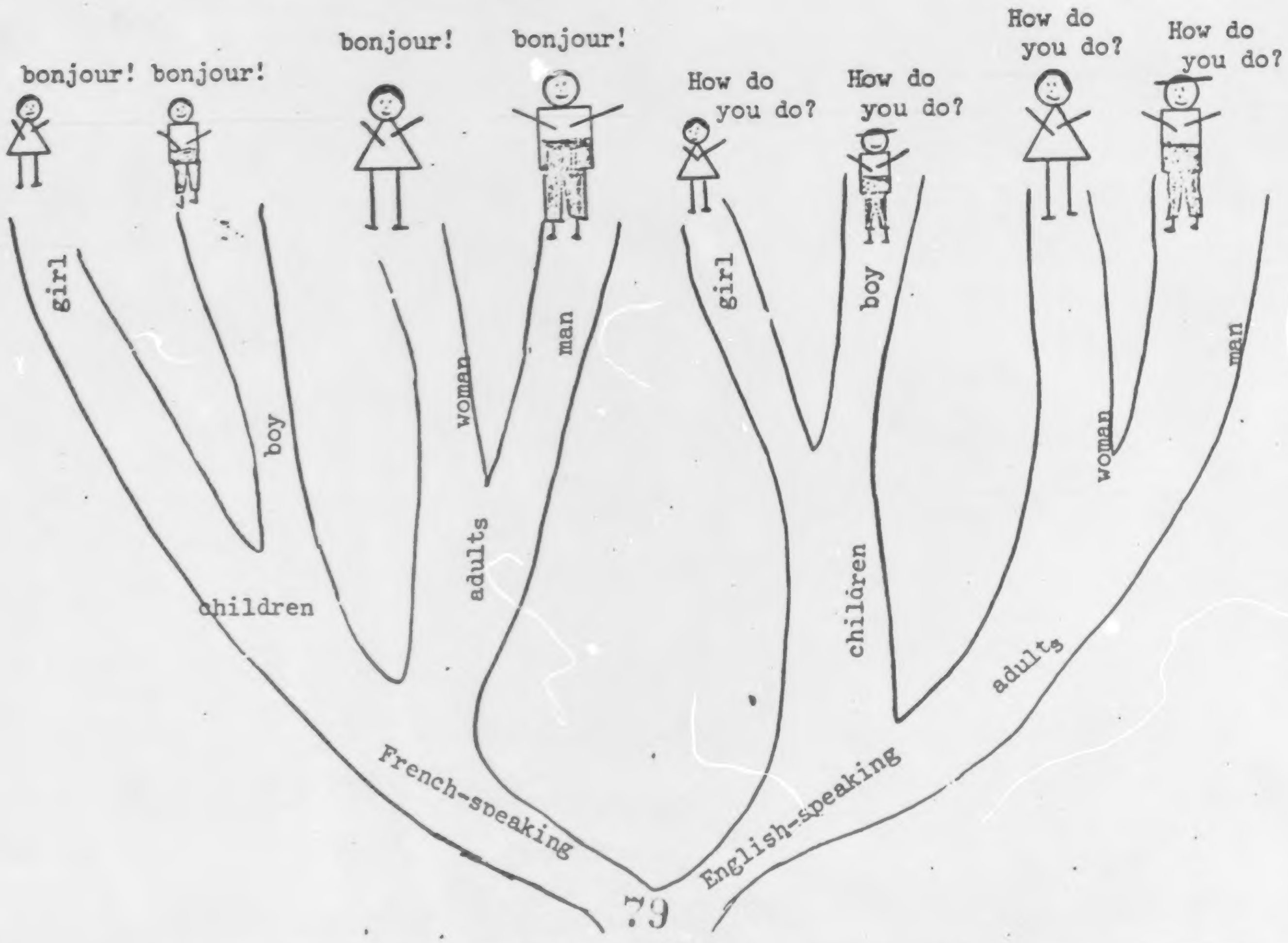

30
15
10
5
6
3
2
1
even
odd
even
odd
even
odd
even
odd
divisible by 3
not divisible by 3
divisible by 3
not divisible by 3
divisible by 5
not divisible by 5

We can see that the non-divisibility by 5 corresponds to paying the fine, non-divisibility by 3 to having mother-in-law over, divisibility by 2 to visiting the friend.

In the same way, we could take eight persons and make a decision. Among these persons, there would be a French-speaking man, a French-speaking woman, and a French-speaking boy, a French-speaking girl also an English-speaking man, an English-speaking woman, an English-speaking boy and an English-speaking girl.

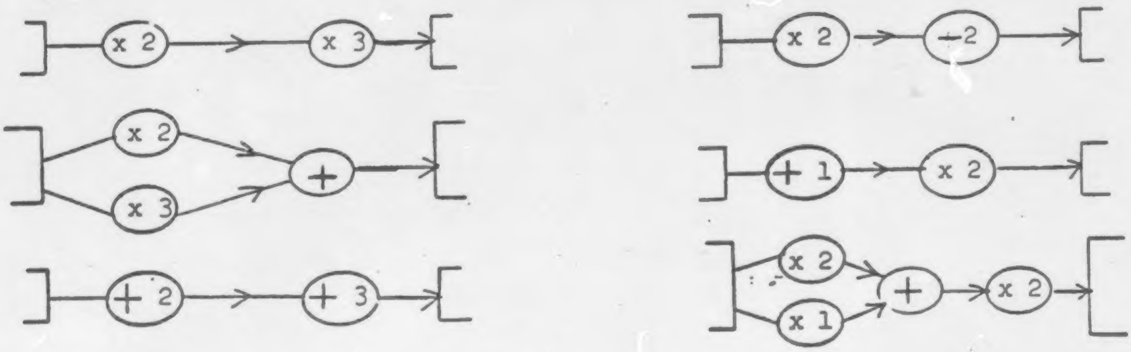

bonjour!
bonjour!
bonjour!
bonjour!
How are you?
How are you?
How are you?
How are you?

Suppose we can pass half-an-hour with one or another of these persons. How are we going to choose between them? If for example, we did not understand children too well, we would undoubtedly choose an adult rather than a child; and on the other hand, if we liked the company of children, we would choose a child; then we might prefer the company of persons of the fair sex to that of persons of the male sex, or the other way round. We must decide if we want to pass half-an-hour speaking French or speaking English. These three decisions being taken we must know which of these three aspects is the most important to us. Is it more important to speak English or French, or to be with a person of the fair sex or of the male sex? Suppose we do not speak English; this means the English-speaking persons are eliminated. Then suppose we prefer the company of children to that of adults. If there are no other stronger reasons to exercise another preference, and we must choose between children and adults, we will choose children. Suppose also that being given the choice between a person of the male sex and of the fair sex, a person of the fair sex will be preferred if there are no other stronger reasons to exercise a contrary preference. This means that the first choice will be a French-speaking girl. If for some reason this French-speaking girl is busy and cannot pass half-an-hour with us, then what is the next choice? The choice will be that of a French-speaking boy. If that boy is busy, then we must choose among the adults, between a French-speaking man and a French-speaking woman. Since we prefer persons of the fair sex, we will choose the woman; if she is busy, we will choose the man. If there are no French-speaking persons at hand, then we will choose an English-speaking girl, then an English-speaking boy, then an English-speaking woman and finally an English-speaking man. It is evident that the structure of this system is identical to that of the two systems already mentioned. Thus, we could use the same tree-structure to represent this third system of choices. Here is this tree-structure.


bonjour!
bonjour!
bonjour!
bonjour!
How do you do?
How do you do?
How do you do?
How do you do?
girl
boy
woman
man
girl
boy
woman
man
children
adults
children
adults
French-speaking
English-speaking

4.5 Arrow diagrams in Arithmetic

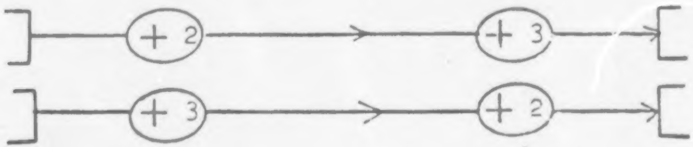
We have already seen that it is possible to use arrow diagrams for arithmetical operators. For multiplication we have already shown it in the description of the factors of natural numbers. It would be possible to do the same for additions and substractions and other arrow diagrams could be constructed where the two kinds of operations are mixed The diagram must show whether the arrow represents a multiplication or an addition. Here are some arrow systems representing sequences of some proposed arithmetical operations.



The space on the left of each diagram is the input to this diagram. Any natural number can be placed at the input. The operations are done according to the indications of the diagram and we obtain the final result, which is the output of the diagram. The output is represented by the space at the right hand end of the diagram.

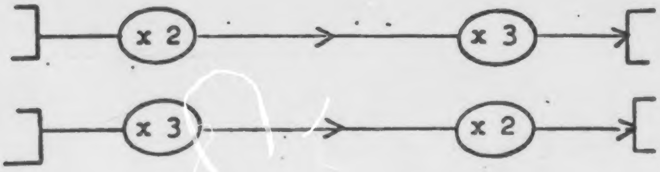
We see that among the arrow diagrams, some are equivalent. Equivalence here means that if we choose the same number for the input of each diagram, we will have the same output for each diagram. If this is always the case, whatever input we might choose, then the two diagrams are said to be equivalent. By using the notion of equivalence between two diagrams, we can introduce notions such as commutativity, associativity and distributivity. Here are the arrow representations for these three sorts of equivalence

(i)



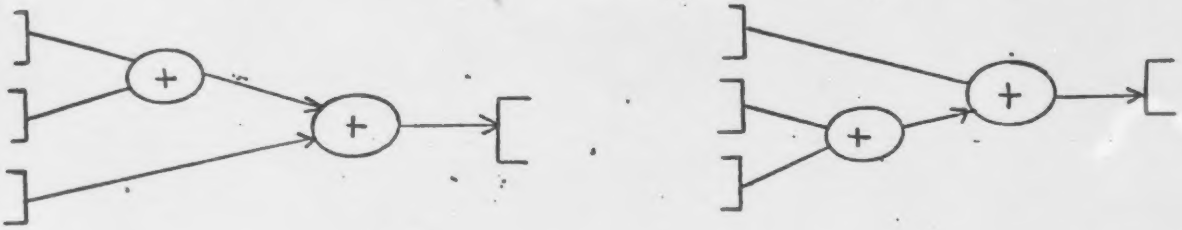
is an example of the commutativity of addition.

(ii)



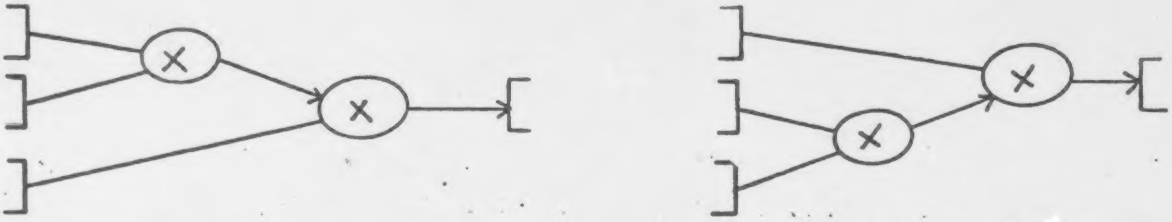
is an example of the commutativity of multiplication.

(iii)



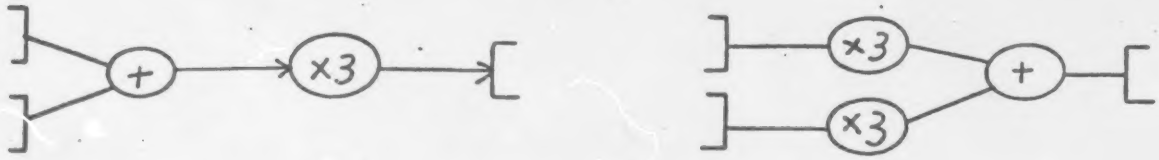
is an example of the associativity of addition.

(iv)



is an example of the associativity of multiplication.

(v)



is an example of the distributivity of multiplication over addition.

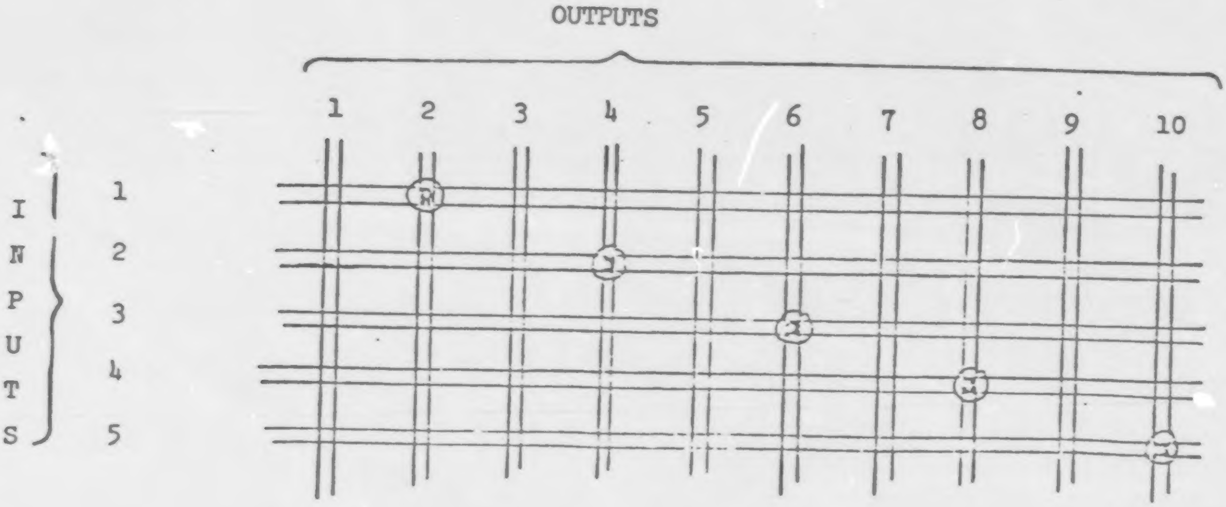
4.6 The Cartesian representation

There are many different ways of representing operators. One of these is the representation called Cartesian. It is done by means of a diagram with a double entry where the values available for the input are to be found at the beginning of one of the rows, and the values we might obtain for the output are to be found at the beginning of one of the columns, or of course the other way 'round! For example, if we want to show the fact that a certain input results in a certain output, we mark the place at the intersection of the row reserved for this input with the column reserved for the corresponding output. Here, for example, are two Cartesian representations. The first one represents the "behavior" of bus number 2 and the other one represents that of bus number 3. The input is the bus stop and the output is the stop where we get out. It is clear that the inputs and outputs in this representation could be replaced by the positions of the regular hexagon which we turned in the second example for illustrating these functions :

We go down

We go up	Albert Street	Frontenac Street	Elizabeth Street	Dorval Street	Downtown	Beaugirard Street
Albert Street				③	$\boxed{2}$	
Frontenac Street					③	$\boxed{2}$
Elizabeth Street	$\boxed{2}$					③
Dorval Street	③	$\boxed{2}$				
Downtown		③	$\boxed{2}$			
Beaugirard Street			③	$\boxed{2}$		

In the case of functions with very few values, arrow representations could be more appropriate. When our universe of elements from which we draw our inputs and outputs becomes numerous enough, the Cartesian representation is better adapted to our needs. Here is the Cartesian representation of the function "twice":



This type of representation introduces the well-know graphs of the functions we see in every school book. In emphasizing this type of representation we would run the risk of giving children the idea that there are no other functions than those which can be represented by means of graphs.

This particular function always generates at the output a number twice the value of the input thus, for example, in the row called 3, there is a sign where this row meets the column called 6; this means that input 3 yields the output 6; the input 2 yields the output 4, and so on. We can see that such a representation can be generalized to take in the rational numbers which fill in the "holes" between the natural numbers: this means that if we took the rational number "half" between 0 and 1, there would be a sign at the place where the horizontal line 1/2 met the vertical line 1 because 1 is twice 1/2. Thus instead of considering a set of isolated points representing the natural numbers appearing in the representation, we will get a set consisting of points densely distributed on a whole line. Again, the first representation shows that our "rows" and "columns" do not have to emanate from points representing numbers at all.

Chapter 5

Building a Mathematics Program

5.1 The Nature of the Problem

The building of a mathematics program must be based on two factors:

1. - The structure of mathematics itself
2.- The development of the child.

The technical problems of the building of a mathematics program consist in reconciling the data of two disciplines: mathematics on the one hand and psychology on the other. There are as yet no programs in existence which are constructed in such a rational manner. The reason for such a lack is probably that mathematicians have not studied the psychological aspects of the problems which must be faced during the learning of mathematics and that psychologists, in turn, have rarely paid much attention to mathematics Those psychologists who are interested in the problem find it a difficult enough task to acquire enough knowledge of mathematics to be able to weigh the learning problems in this discipline adequately, so as to formulate hypotheses which can be checked by experiments in the psychology laboratory. Mathematicians have for the most part contented themselves in writing texts and have not imagined that attention to the learning of the contents was their responsibility; in their eyes this responsibility was entirely the teacher's.

During these last years, a certain amount of work has been done by different research teams, particularly those related to the International Study Group for Mathematics Learning.. Naturally, the solution to the problem of program-construction cannot be unique. There must be a great number of solutions which reconcile the data of the psychology of learning with those of the structure of mathematics itself. We have to admit that the proposed solution is only one among many possible ones; yet we believe it to be better adapted to the demands of the modern world than the usual traditional or "modern" mathematics programs with their accompanying methodologies.

5.2 Sets

In general mathematicians observe the formalist point of view, that is, they suppose that mathematics is based on logical and set-theoretical principles, and that it is possible to build on these the rest of mathematics by taking into account the relations between logical operators or other operators. Consequently, it is a priori probable that the learning of properties of sets should precede the learning of the properties of numbers. We can go back further and ask the following question: "What must we know about sets?" Naturally, we must know what their elements are, where we can find "candidates" that can "be an element of a given set". In the child's experience, the elements of sets are often events or preferably objects. How can we decide if an object belongs or does not belong to a set? To do this, we must be able to distinguish one object from another. This requirement means that we must have knowledge of the properties of objects and so we should be able to classify them. It seems that our first step in planning the understanding of natural number is to provide exercises for classifying objects into categories. This means arranging various objects in order and in different groups. This process leads us to us to the establishment of criteria which allow us to decide if an element belongs or does not belong to a given set. Classifications, followed by double classifications, will lead us to conjunctions and disjunctions. Establishing classifications of objects leads us to recognize that some are of one kind but others are of another kind. Within a certain set, all objects are either of one kind or of another kind. This leads us to exclusive disjunctions. Later on we will introduce other classifications where there will be overlapping. For example, classification might be done by color and by shape; children are classified according to their sex, the color of their shoes or of their clothes. If we unite the two "mixtures" there will be overlapping. There is no overlapping if we classify children according to their sex:boys or girls; but if we classify children not only by their sex but also by the color of their shoes, whether black, red or of another color, there will be overlapping because there will be girls as well as boys wearing black shoes; and there will also be girls as well as boys wearing shoes which are not black.

It is fundamental for all of us to learn to distinguish objects depending on whether they possess or do not posses a certain property. This distinction can lead us to the idea of belonging to a set. When we have learned the meaning of one property, we shall be able to proceed to learning how to handle a number of different properties. This is what we do when double-entry classifications are used. This leads us, from a logical point of view, to conjunctions and, from the point of view of sets, to intersections. When we examine a set in which all the elements possess one kind or another kind of property, we have a disjunctive classification; from the point of view of sets, we proceed to the union of the set of elements possessing the first kind of property with the set of elements possessing the other kind. Consequently, the conjunction of attributes leads us to the idea of intersection and disjunction of attributes to the union of sets.

Very early in the learning process, negation will have to be introduced. This is necessary, as much to develop the child's methods of thinking as to introduce its analogue in arithmetic, that is, subtraction. When a set is subdivided into two complementary sets, the result is a situation in which each element of that set belongs to one or the other of the two complementary sets but never to both at the same time. Consequently, we can say that if an element does not belong to the first sub-set, it will belong to the second, and if it does not belong to the second then it will belong to the first. Later on, when we come to subtractions, we will have to introduce the set operation known as the difference between a set and a subset of this set. We subdivide a set into two complementary sets and we remove one or the other of these sets. The operation consists in finding what is left, that is, the other sub-set. We will have to distinguish between this operation, namely that of finding the difference set and that of subtraction in the case of numbers.

To introduce addition, we have to resort to the operation of the union of two sets. We must distinguish two kinds of unions; the first kind , in which the sets united can have common elements; and the second kind, in which the sets united do not have common elements. In this last case, the intersection of the two sets united must be empty. Evidently, there is here another notion to be introduced, namely that of the empty set and its analogue in arithmetic, namely zero. The empty set is the set in which there is no element and zero is the number of elements in that set. They must not be confused with each other.

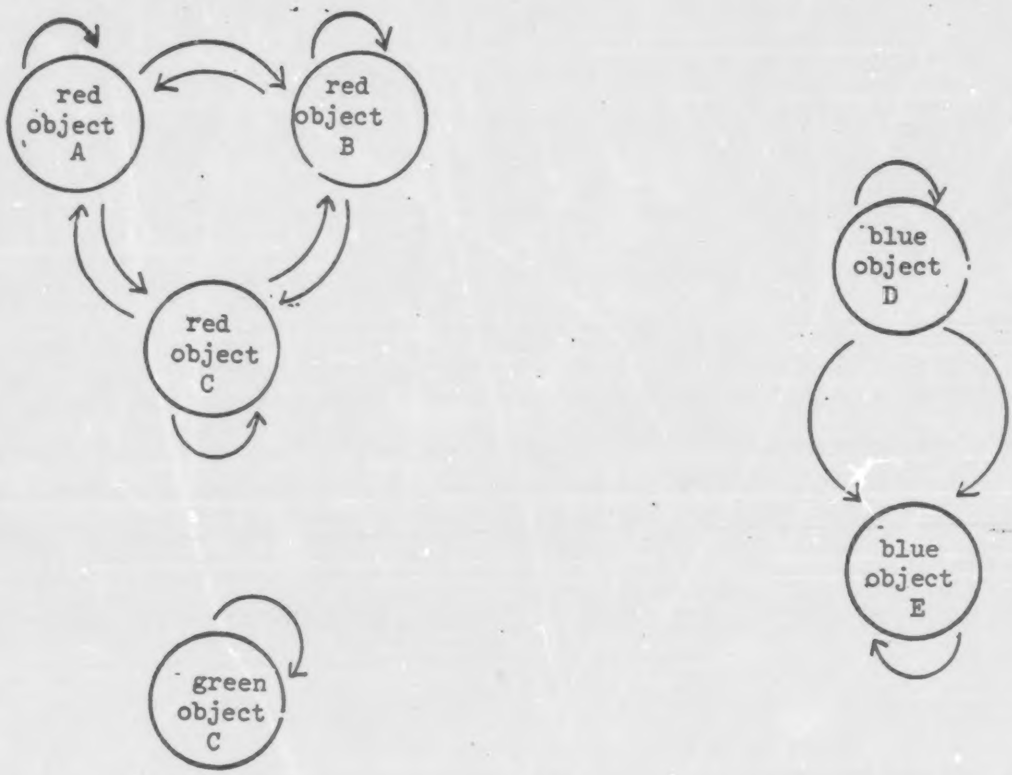
The operations on sets described above allow us to produce other sets every time we have two sets available. There is also an operation called "taking the complement"; this allows us to build a set when we have only one set available. An operation for which two entities are needed to find the resulting entity, is called a binary operation; an operation which calls for the use of only one entity to find the resulting entity is called a unary operation. Finding the intersection of two sets, finding the union of two sets, are examples of binary operations; finding the complementary set of a given set is a unary operation.

5.3 Relations

In our attemps to construct a program, we shall not only be concerned with operations on sets, but also with operations on the elements of such sets. To do this, we will have to link some elements to other elements. This is a necessary part of our requirements leading us to the idea of natural number. Linking one thing to another always requires two things. In mathematics, we use few relations but it will be necessary to know how to handle the most important ones and to get to know their properties well before we can use them efficiently.

The most important relations in mathematics are those of equivalence and of order. There are also the relations of difference, logical relations, etc ...

Equivalence relations can be found linking objects or linking sets of objects. For example, if in a given set of objects, each element is of one particular color, the sentence: "This object is of the same color as this other object" becomes an equivalence relation.



Here is a representation of an equivalence relation. The arrow reads: "is the same color as". For example, we could read:

object D is the same color as object E

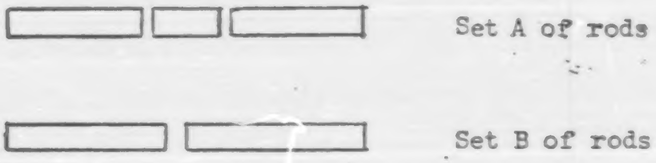
or

object C is the same color as object B,

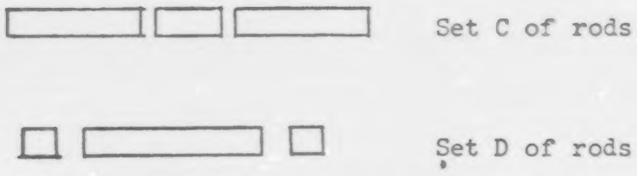
but we could not read:

object E is the same color as object C, since there is no arrow going from E to C. An equivalence relation subdivides our world of objects into disjoint classes; they are called equivalence classes. All the relations of resemblance which produce such a subdivision into disjoint sets are called equivalence relations. Naturally, we can talk of equivalence relations between sets. One of these relations, introduced in modern mathematical workbooks, is the relation of equipollence, that is, "to have the same number of elements as". Many children believe this to be the only kind of equivalence existing. When they get to the study of fractions or of matrices and find equivalences again, they can not imagine that this kind of equivalence is really the same as that of equipollence between given sets. It is evident that it is difficult to learn what an equivalence relation is if only one example of this notion has been provided!

The use of concrete materials presents an opportunity as well as a danger in the introduction of equivalence relations. The use of Cuisenaire rods gets us used to the idea that there is only one kind of equivalence: the one which allows us to say that this set of rods takes up the same space as that other set. Evidently, we should also get familiar with the equivalence relations which tells us that "this set of rods possesses the same number of rods as this other one". "I have the same number of rods as you have" expresses an equivalence relation just as "I have the same quantity of wood as you have" which expresses another equivalence. If the child has played with a good number of different equivalence relations, as well as with other relations between objects and between sets, he will be able to learn the properties of the important equivalence relations that he must learn to handle.

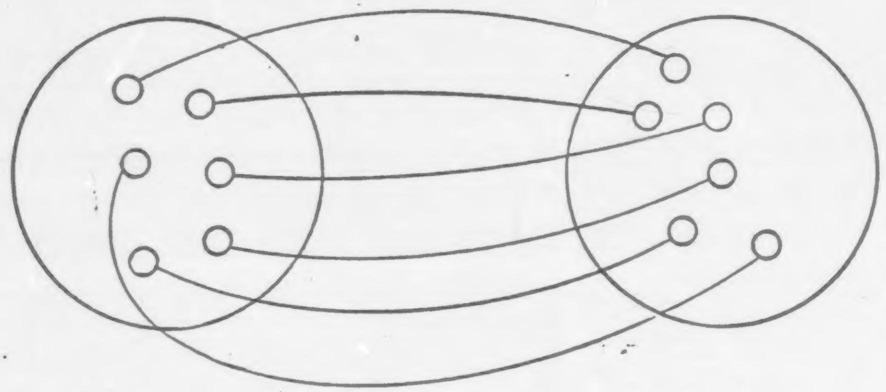


It is true that: Set A "occupies as much space" as set B.
It is not true that set A "has as many elements as" set B.



It is true that set C "has as many elements as" set D.
It is not true that set C "occupies as much space as" set D.

For example, we can say that there is an equivalence relation between sets, if we can establish a one to one correspondance between the elements of one set and the elements of another. This criterion is to be used for the kind of equivalence called equipollence. In common parlance, this equivalence links one set to another if the first one has exactly as many elements as the other.



A "has as many elements as" B. A and B belong to the same family of sets. In mathematics, they are said to belong to the same equivalence class. The name of this class is the cardinal six.

We shall see in the next section how cardinals and ordinals must be combined for the construction of what is called the natural number.

5.4 The natural Numbers by the Synthesis of Equivalence and Order

The application of an equivalence relation which says: "This set possesses as many elements as the other" clearly subdivides the world of sets into disjoint classes. These classes or families are called natural numbers. There are those who prefer to say that it is the properties of these sets belonging to these families which are called the natural numbers. Property three is a property of all sets which possess only three elements, no more, no less. This property tells us that it concerns that class and not another class where there can be less than three elements or more than three elements. There is no overlapping of these classes, because in the same set there cannot possibly be three and four elements at the same time.

After having established this set of families, the families thus produced can be put in order. We can invent the "more" and "less" game. Given two families, we can take a set from one family and a set from another family; we can compare them element by element, that is, all the elements of one set will be compared with all the elements of the other. If in one set, there are elements we cannot associate with any elements of the other, then, it is said that in this first set, there are more elements than in the other and that the family to which the first set belongs is larger than the family to which the second set belongs. In this way, we have found a means of deciding, given two families, which is the larger and which is the smaller of the two. We say that if family A is larger than family B, then family B is smaller than family A.

We can now introduce the "one more" game. One set will have one more element than another, if, and only if, after the association of all possible pairs, there is only one element from the first set which can not be associated with any element of the other set. In this case the family, from which we have taken our first example, will be "one more" than the family from which we have taken the second one. This idea of "one more" establishes a chain which begins with the family called "zero", and the members of the chain will be the successive cardinals, one two three, etc. Each cardinal is joined to the next by the "one more machine".

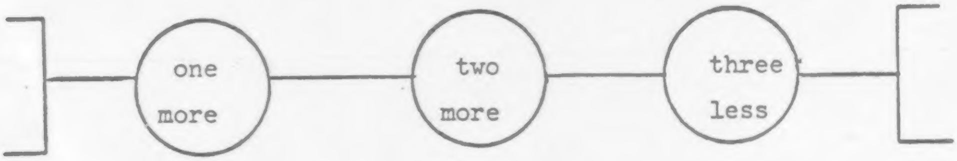
The element we choose, to put into the machine, is called the input and what is done to change the input into the output is called a function or operator. The element which the machine produces is called the output. Our machine is automatic; it can produce only one possible output as soon as the input or inputs are chosen. Many machines have more than one input, for example the binary operators. We also might have to do more than one thing before the output is reached, but the latter is always fixed once the input or inputs are chosen, otherwise our relation would not be called a function or an operator.

To sum up, the consideration of relations and of sets in general leads to the consideration of equivalence and order relations in particular. Thus we realize that equivalence classes are produced by an equivalence relation. The relation of "having as many elements as " is a very important equivalence relation linking sets. The equivalence classes or families generated are the natural numbers one, two three.... . They are arranged in order by the relation "more, less". The "one more" relation finally determines that each family produces a next family. This family is the one which immediately follows a specific given family.

5.5 Other Roads leading from Relations to Functions

At this point, we must seriously think of the psychology of the process we are examining. By taking a narrow enough road, we have come to the idea of one particular function. We have followed this arithmetical road by considering sets and relations. There are many possible and different relations which will lead us to other possible operators, to other functions, for example, of a geometrical nature. The exploration of the simple properties of space can be undertaken with children in kindergarten or in first year of elementary school. We can begin with exercises which will lead them to the appreciation of geometrical relations and later to geometrical operators. For example, a game could start as follows: the teacher will start walking and will say: "I am walking between the blackboard and the bookshelf"; the board will be on one side of the class and the bookshelf on the other. Suppose the door and the window are on the two other walls; then the teacher can say: "I am walking between the board and the shelf but you are walking between the door and the windows. What must I do to move as you are doing?". The answer is simple: the teacher must make a right-angle turn. It does not matter in which direction he turns; whether he turns right or left, he will end up walking between the window and the door and will do as the child who is walking between these two. If "as you" means he is walking not only between the window and the door but that he is also walking between these two walls and in the same sense . as the child, that is, toward the wall, the teacher has no more choice, he must turn about in one particular and specific sense, that is, left or right. In this case, the notion of "at a right angle", which was a general relation in the first case, becomes an operator which says: Make one quarter of a turn to your right or else make one quarter of a turn to your left. In this case the result of the action is well determined whereas "take a right angle turn" has no specific result because we can turn to our left or to our right if the sense has not been specified.

There are a number of means of tracing the road leading from relations to functions; concrete material such as logic blocks can be used. But let us now continue our discussion on functions. Naturally the "one more" operator will not be left alone; we will soon introduce the "one less" operator, the "two more", the "three more", operators, etc ... and the children will link these operators with one another. They will take a certain state of things on which they will operate and another state of things will result. They will operate on the new state and will obtain still another state after having operated again by means of yet another operator. Thus they will be able to build chains of operators such as for example



There will be states between the operators. As soon as the first state is chosen, the others will be determined.

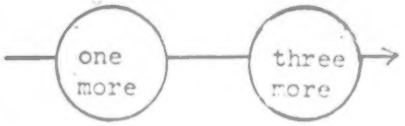
The states in arithmetic are expressed by means of natural numbers but they can be represented in a concrete manner by a set belonging to the family concerned. For example if we want to represent the natural number 3, we must take a representative sample of the family 3. This means any set of three elements. Now if we want to operate by the operator "two more", we must find another representative set of the family three, unite with it a set of two elements which has no common elements with our representative set of three elements we have chosen, and determine the family membership of this union-set. This family will be the family five. This end state will be represented in a concrete way by the presence of a set of five objects.

At the beginning, it will be difficult for the children to use more than one operator after another. For example, if they are asked to operate by means of the operator "one more" and then to operate on the state obtained once more by the operator "one more", they will not necessarily know that they could have effected the total change by taking the operator "two more". In other words, they will not know that the operator "one more" followed by another operator "one more" could have been replaced by "two more". More precisely, the child might know that state *one* operated on by the operator "one more" leads to state *two*. He might not know that the operator "one more" followed by an operator "one more" can be replaced by the operator "two more".

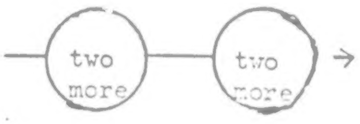
But careful! What is meant by "*could be replaced*"? Here another equivalence relation is being introduced! We know that an equivalence relation must always operate in a specific universe, in a universal set of certain objects or entities which may even just be imagined in our mind. What are these "objects" we want to link to others of its kind so as to build equivalence classes? These "objects" are called *chains of operators*. The simplest chain consists in one operator only. Chains can become more and more complicated, as they can consist of operators applied one after the other always using the output of the last operator as the input of the next operator. What is meant by saying that a chain is equivalent to another?

It surely means that the input - output property of one chain is identical to the input - output property of the other. If for example an input 1 yields an output 5 in the case of one chain, then input 1 must yield the output 5 for the other chain. This must be so for any given input. Two equal inputs must yield two equal outputs in each chain. If two chains possess the same input-output property, these chains belong to the same family of chains, to the same *equivalence class*.

For example the chain



is equivalent to



5.6 Operators become States; Introduction to Integers

Here we come across a great psychological difficulty. It is hard to talk of operators or of chains of operators as entities which we may transform into other entities. At the beginning, "one more" was an action, to be performed, when a state was given and from this action another state had to result. Now "one more" becomes a chain which we can link by means of an equivalence relation to another chain. This means that what was an action now becomes a thing. In other words, the operator has become a state. Every time an action becomes a thing in our mental constructions, we pass on to a higher level of abstraction, because we operate on an operator. In terms of the epistemological theory developed by Piaget's school, this means that in this particular area we have reached the operational stage. In the pre-operational stage we could only operate on states and so reach other states. When we have reached the operational stage, we can operate on an operator and obtain another operator. The difficulty which has to be faced before taking such a step must not be underestimated. For that reason, states and operators are studied as separate entities for some time in our program; we try to make this separation in every area of mathematics whenever a start is made to study that area.Each time a new structure is introduced, we make a distinction between states and operators. For example when we study the geometrical properties of the square, of the equilateral triangle, of the regular tetrahedron, we take the possible positions in which such a solid or such a plane figure can be placed; these are called states. The movements which can be effected to change the solid or the plane figure from one position to another are called operators. When children are still at the pre-operational stage, they learn how to make up chains of states and operators which follow one another, particular positions are given which the children change by means of particular operators which lead to other particular positions. These positions will be changed again by applying further operators which will lead ultimately to other states. The kinds of problem which can be solved at this stage are for example:

(i) What will be the state which will result from a particular state if a particular operator is applied?

(ii) What operator should be applied to a certain state to obtain a specific final state?

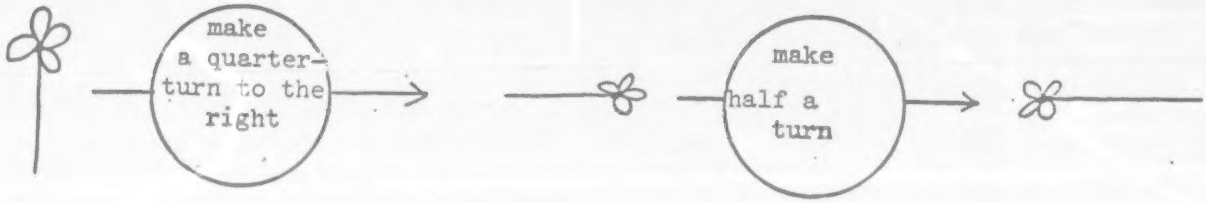
When the child has reached the operational stage, he will be able to consider the movements themselves as entities which he will transform, and he will be able to solve problems such as:

Given a series of operators which make a certain change of state, how could we replace this series of operators by another series of operators which would produce the same change of states?

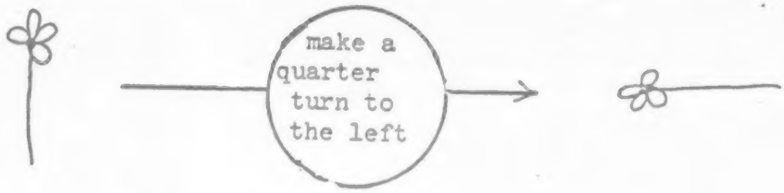
Eventually the child must realize that if the two series behave in the way as far as changes of states are concerned, they will be equivalent. These operators can be additions and subtractions, but also rotations, reflections or other changes whose results are determined once an input has been chosen.

Here are certain examples:

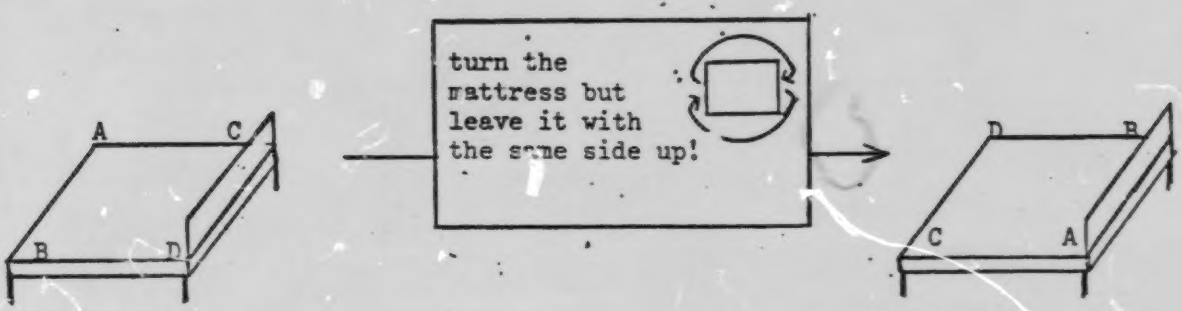
(i)



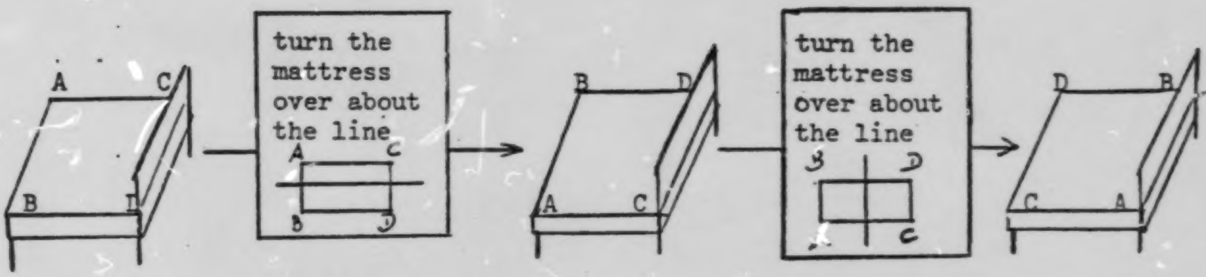
(ii)



(iii)


A
C
B
D
turn the mattress but leave it with the same side up!
D
B
C
A

(iv)


A
C
B
D
turn the mattress over about the line
A
C
B
D
B
D
A
C
turn the mattress over about the line
B
D
A
C
D
B
C
A

There are only four possible positions for the mattress. The reader could try to identify them and arrange for the four positions to follow one another regularly. We can do so by using only two kinds of movements.

Let us now examine more closely how series of operators can act as states and how such series can be transformed into other series. We have seen that a numerical state can be described in a practical way by the presence of a particular set belonging to a family we want to represent; for example, if we think of state 3, we can represent that state concretely by providing an element of that family, that is, any set possessing three elements. Now we must provide a concrete example of the operator "2 more" and one way of doing it would be to give one input and one output. For example, we could take three apples and five oranges; there are two oranges more than there are apples, and thus our chosen pair of sets could represent the property 2 more. We start with the set of apples; we wish to arrive at the set of oranges, the number of apples being 3, the operator "2 more" must be applied in order to get the number of oranges, which is 5. We know that a set of three elements can be expressed by any set possessing three elements; in a similar manner, here we can represent the property "2 more" in many different ways. If we take three apples and five oranges or four apples and six oranges, or ten apples and twelve oranges, we will always have 2 more oranges than we have apples. Consequently, all these representations would be convenient aa a concrete realization of the property "2 more". Naturally, we can use girls, boys, knives, forks, or any two kinds of objects which can be easily distinguished from one another. In this way, by providing an input set and the corresponding output set, we can give a great number of concrete representations of the abstract notion of having a certain number more of one sort of thing than of another sort.

Representations "having less" can be provided, as well as of "having more". It will even be possible to combine these properties by operations. For example, we will be able to associate the property "to have three less" with the property "to have two more" by taking a representation of the first property and a representation of the second one. Suppose that the property "to have two more" be represented by the union of a set of ten apples with a set of twelve oranges. The property "three less" could be represented by the union of a set of ten apples with a set of seven oranges. If we take the union of these two sets, we will get twenty apples and nineteen oranges; we will then have one orange less than we have apples. In other words, we will have obtained a representative of the property "to have one less".

Thus we are in a position to introduce the addition and the subtraction of integers. For example, subtraction could be embodied as follows: let us take the same properties, a set of apples and of oranges representing the property "two more", for example a set of ten apples together with a set of twelve oranges and then let us try to remove a set representing the property "three less". Such a set could be the union of a set of eight apples with a set of five oranges. Five oranges are "three less" than eight apples. To do the "subtraction" we take away the union of the two sets representative of the property "three less", from the union of two sets representative of the property "two more". We get two apples and seven oranges. This means we have obtained five oranges more than apples. Consequently, "two more" subtract "three less", results in "five more". We must realize that any set representative of the first property and any other representative of the second one would have served equally well and that the final result would have been the same as regards the number property of the final set, we would always have obtained the property "five more", that is, five more oranges than apples. Thus we can build up many concrete situations which help us to embody the addition and the subtraction of integers.

5.7 The Introduction of Rationals

We have seen that being in a position to consider "one more", "two more", "one less", "two less", etc ... as states, has helped us to introduce integers, together with addition and subtraction of these numbers. In the same way, we could use multiplicative operators, that is, the operations of multiplication and of division.

One way of introducing multiplication and division is to study them as replacement operators. The operator replaces every element of the input by a certain set of elements at the output. All sets replacing input elements at the output must belong to the same family of natural numbers. For example, each element of the input set could be replaced by a set of three elements , which would be united and would make up the output-set. The number of elements of this set at the output would be the output number. The number of elements of the input-set would be the input number. For example, if we are given a number of plates and on each plate we place three oranges, we have introduced the operation which provides us with three oranges for one plate. If we have taken four plates, the input number is four. A set of 12 oranges will be our output-set, because we will have "replaced" each plate by 3 oranges.

The division operators can be introduced in the same way. They are the inverse operators of multiplications. Suppose that there is room in one car for four passengers; if the would-be passengers wait in line and we notice that there are twenty of them waiting, we know that we will need five cars to transport them. Each time we take a set of four passengers, we can fill up one car. Here the input-set is a set of twenty passengers and the output-set is a set of five cars. The operator is one car for each set of four passengers. By following this type of approach, children will soon get the idea of the fractional operators, because there is no reason why we should not take "many for many" operators. For example, we have a certain number of chairs for seating children wanting to watch a television program. We notice that there are too many children. We might say "we can put two chairs for three children and they must do their best to find a place to sit on. If twelve children are present; and we want to seat them, we place two chairs for every set of three children. Thus we will need eight chairs for the twelve children. Here the input-set is the set of twelve children and the output-set is the set of eight chairs. The requirement that tells us to find a set of two chairs for each set of three children is the operator. This requirement can be numerically expressed as the operator "two for three". We say that the number of elements in the output-set is 2/3 of the number of elements of the input-set. The symbol 2/3 will be considered as the abbreviation of "two for three".

Up to now fractional operators have been actions. If we want operators to become states, we must provide examples of inputs and of corresponding outputs, as for example the set of twelve children together with the set of eight chairs. We could also have taken six children and four chairs which would have led us from the set of children to the set of chairs. Here we are establishing equivalence classes. These equivalence classes will become classes to which the fractions belong which are said to be equivalent. Children who have studied equivalence relations and have realized the way in which an equivalence relation subdivides the universe into disjoint classes (equivalence classes), will have no difficulty when faced with the problems of the equivalence of fractions. They will see the resemblance between the pair of sets the first of which is a set of twelve children and the second, a set of eight chairs and the pair of sets the first of which is a set of six children and the other a set of four chairs, because in both cases there are always two chairs each time there are three children. There is an infinite variety which can be used to represent the fraction 2/3 (two for three). We can use chairs and children, apples and oranges, stones and shells or any other objects which can be easily distinguished from others and for which a suitable cover story can be invented to amuse the children.

To resume: We can provide concrete experiences for the introduction of natural numbers by means of sets and of equivalence and order relations which lead to the understanding of additive and later of multiplicative operators. Moreover, the possibility of using operators as states can lead from the study of additive operators to the study of integers on one hand, and from the study of multiplicative operators to the study of rational numbers on the other hand. Evidently, addition and subtraction are operations easy to introduce for the integers, and multiplication and division are relatively easy to introduce for the rationals. The addition of integers can be done through putting operators end to end, taking only additive operators; multiplication, in the case of rationals, can be introduced by a similar process, taking only multiplicative operators. We shall encounter more difficulty in the case of integers when multiplication is tackled. In the case of rationals, addition and subtraction will present the greater difficulties.

5.8 The Klein group and the cyclic group with four elements

To tackle these problems, it will be advisable to introduce very simple group structures. For example: the cyclic group with four elements and the other group with four elements, usually called the Klein group, can be presented in a very elementary way so as to give children the means of tackling the problems of multiplication in the case of the integers.

Each time we consider two variables each of which can only assume two values, we have a representation of the Klein group; but each time we consider only one variable which can assume four values constituting a cycle, we have a representation of the cyclic group of four elements. For example we can play a game with sets of boys and girls. If we accept the empty set as legitimate, there are four ways of making up sets of boys and girls:

i) empty set, with no boys, no girls

ii) set where there are only boys

iii) set where there are only girls

iv) set where there are both boys and girls.

We can introduce changes which will bring us from one case to another. For example:

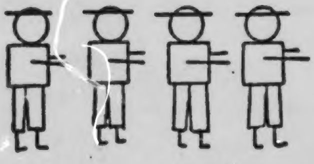
changes in girls means: if there are any girls, we sent them away, or, if there are none, we bring some in;

change in boys means: if there are any boys, we send them away, or, if there are none, we bring some in;

change in boys and girls means: we change the presence or the absence of boys as well as the presence and the absence of girls;

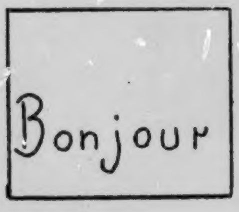
no change means: we leave the situation as it is.

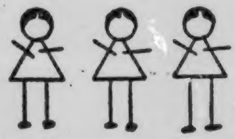
We can compare these transformations to the geometrical transformations of a rectangle which must always occupy a certain space. The figures below will help the reader to make the necessary comparisons.



sample of

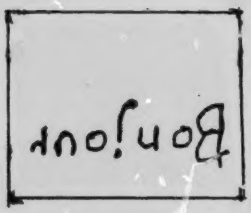
" there are only boys"

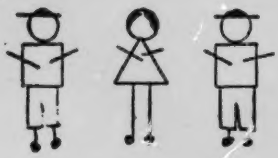




sample of

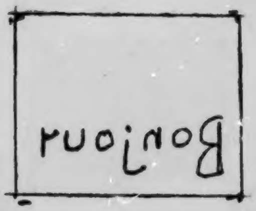
"there are only girls"





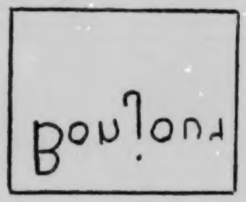
Sample of

"there are girls and boys"



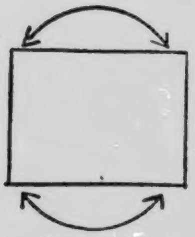
Sample (the only one possible)

of "there are neither boys nor girls"

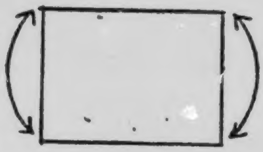


Let us write something on a transparent rectangular sheet of paper; our rectangle can be placed in four different ways so as to occupy a certain space.

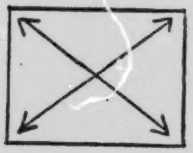
To make the changes from one position to another, the following movements may be made



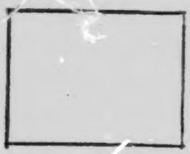
Turn like pages of a book



Turn like pages of a calendar



Half-turn without turning leaf completely



Do nothing

Let us take some more examples. Consider the passing of time: we speak about things happening a quarter of an hour later, half-an-hour later, a quarter of an hour before or "now". Let us take our "now" at a time "on the hour". The passing of a quarter of an hour brings us from one state to the next. If we take the passing of a quarter of an hour twice, we get two quarters of an hour, that is, half an hour. The passing of three quarters of an hour amounts to looking at the time a quarter of an hour ago, as far as the minutes are concerned. Thus the quarter hours provide a simple means of creating a cycle of four. Equally, the quarter turns we use when turning about a vertical axis about our own body, leads us to situations where the cyclic group of four elements is used. The movements of a rectangle or the bringing in and the sending away of boys or girls in a set lead us to the manipulation of the elements of the Klein group. Here are other examples of the Klein group and of the cyclic group of four elements.

(a) Here are four sentences:

You come to school P You do not come to school $\overline{P}$

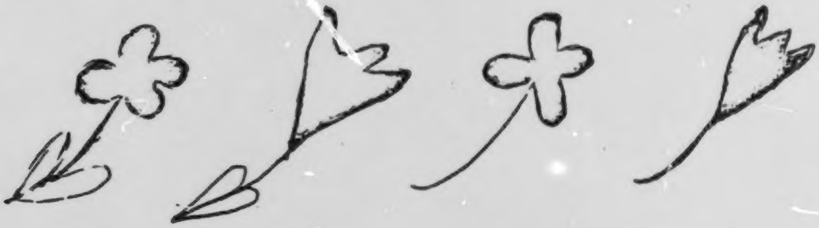
Do you come to school P'? Do you not come to school? $\overline{P}'$

We go from P to P' by changing the affirmative form to the interrogative form. The change $P \longrightarrow P'$ or $P' \longrightarrow P$ is a passage between these two forms of the sentence, in one or the other of the two directions possible.

The passage between P and $\overline{P}$ is a change between the positive and the negative form.

Evidently the two aspects of the sentence can be changed at the same time, for example between P and $\overline{P}$. This example illustrates once more the structure of the Klein group.

(b) Here again is a concrete example. Take the four flowers:



we can go from one flower to another

i) by going to another kind, but without changing the leaves

ii) by going to the same kind, but with a change in the leaves

iii) by going to another kind with a change in the leaves

iv) by keeping the same flower.

It is easy to see that we have embodied the same mathematical structure once again.

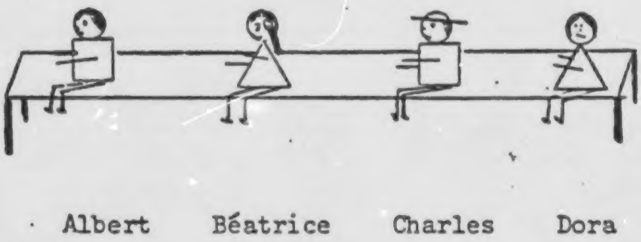
(c) An arithmetical example would be multiplication modulo 8 of the numbers 1, 3, 5 and 7. The expression "modulo 8" means that we divide any number we might obtain by 8 and consider only the remainder as our "answer". For example:

$3 \times 5 = 15 = 8 + 7 = 7 \bmod 8$

$5 \times 5 = 25 = 24 + 1 = 1 \bmod 8$

The numbers 3, 5, 7 play the parts of changing operators and number 1 plays the part of the neutral element.

(d) A 'cyclic' game would be the following. Seat four children in a row:



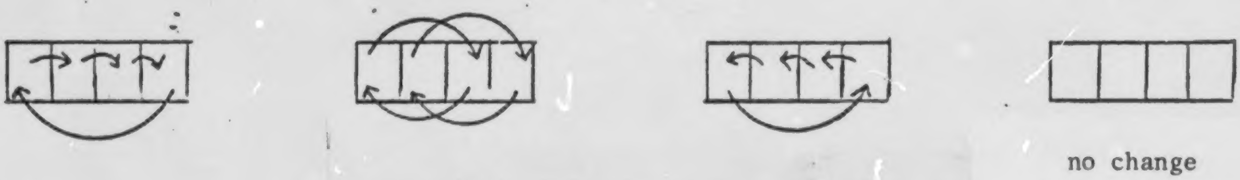
A change to be made in order to create other distributions of children would be that the child on the right take the place of the child on the left and that the others move up one space to the right. If we start from Albert, Beatrice, Charles, Dora, we get the following row:

Dora, Albert, Béatrice, Charles

the four possible rows are

ABCD, DABC, CDAB, BCDA

the changes are the following



This then is quite different in structure from the other three examples, which represent the Klein group.

5.9 The Structure of the Ring

Let us now use the boys and girls in another way. Let us distinguish between tall ones and small ones, as well as between boys and girls. A set of children will be classified according to two criteria:

(i) the difference between the number of small boys and small girls,

(ii) the difference between the number of tall boys and tall girls.

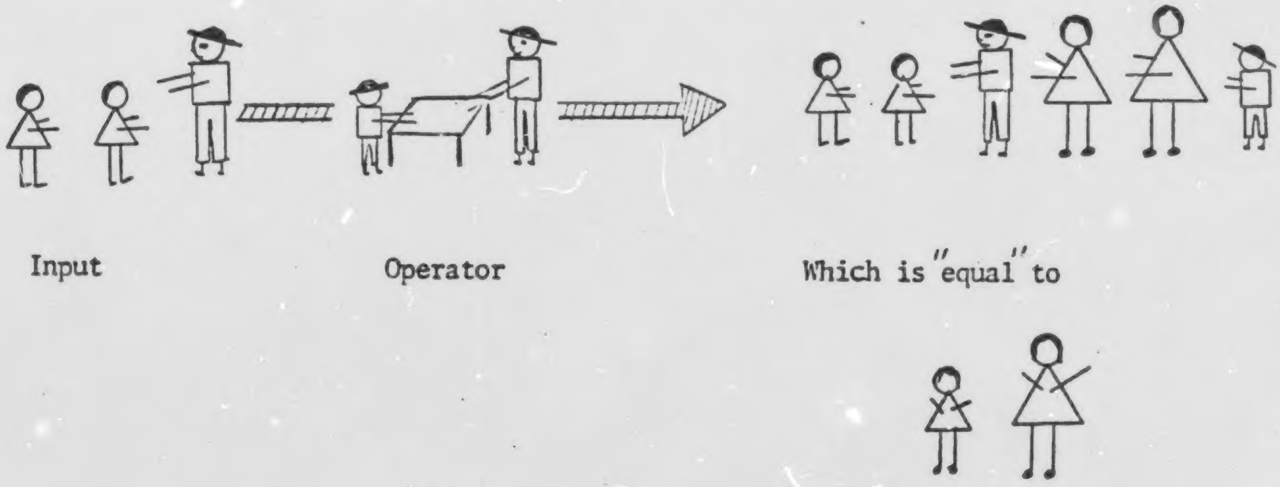
This difference will be a positive (negative) integer if the boys(girls) outnumber the girls (boys).

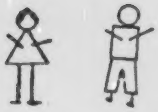
Thus we will need two integers to specify the class to which any given set of boys and girls belongs. (2,-1), for example, would specify the class of sets in which the *small* boys outnumbered the *small* girls by two, and the *tall* girls outnumbered the *tall* boys by one. Addition of such pairs would follow the ordinary rules of addition for each integer in the pair.

To introduce a second operation we could play a game in which one set of children was replaced by another as follows: The neutral element of the game would be to replace each small boy by another small boy, each small girl by another small girl, each tall boy by another tall boy, and each tall girl by another tall girl - that is simply to leave the team as it is. We could represent such a replacement by the presence of a *small boy* , who would act as the "neutral" operator. If a small boy knowked on the table this would mean no change. On the other hand, if a small girl knowked on the table this would (by our stipulation) mean that the small boys must be replaced by small girls, the small girls by small boys, the tall boys by tall girls and the tall girls by tall boys. What would happen if a tall boy knocked on the table or if a tall girl knocked on the table? Again we could invent any rules we liked. For example, if a tall boy knocked on the table the small boys could be replaced by tall boys, the tall boys by small bc the small girls by tall girls, and the tall girls by small girls. Finally, if a tall girl knocked on the table, the small boys could be replaced by the tall girls and conversely, the small girls by the tall boys and conversely.

We have specified a set of rules for "replacements". The children will have already learned that replacement operators can act as multiplicative operators, thus they will realize that this is an operation which could serve as a multiplication. The operation which consists in putting together two sets of children, will act as addition. The elements on which we operate are ordered pairs of integers. Multiplication is defined through the transformations effected by a small boy, a small girl, a tall boy, a tall girl. If we "multiply" by more than one person, the rules agreed upon for each of these persons must be respected. For example, if one small boy and one tall boy knock on the table at once, this means that each person is replaced by two persons. The first person to replace the person leaving follows the rule agreed upon for the small boy 'nocks' and the second person to replace the one leaving follows the rule of the tall boy. The following sketch shows the transformation:

tion:



It can be verified that if  are taken for our input, we get a set which is described by zero small ones and zero tall ones. We should remember that zero here does not mean "NONE", but that it means "as many boys as girls".

The reader will have noticed that the replacement game follows the rules of the Klein group. For example, let us take an input set

(3 small boys, 2 tall girls)

(Symbolically this set could be represented by the pair (3,-2))

A small girl knocks on the table to give the output

(3 small girls, 2 tall boys)

and then

a tall boy knocks on the table to give the output

(2 small boys, 3 tall girls).

We could have gone directly from the input set

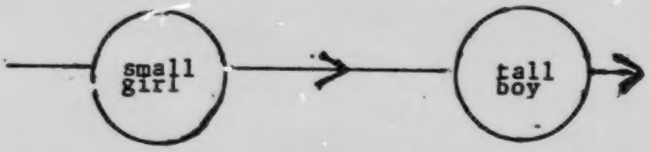
(3 small boys, 2 tall girls)

to the final output

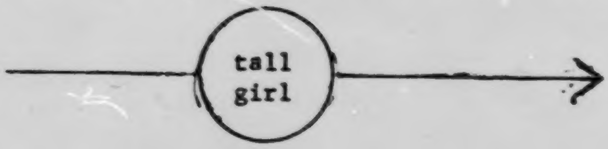
(2 small boys, 3 tall girls)

if a tall girl had knocked on the table

Thus the chain of operators



is equivalent to the "chain" consisting of the single operator



It can easily be verified that the table for combining the four operators which consist of one person only has the same structure as the table for combining the operators which turn the mattress over (see page 93). This is the structure of the Klein group.

The chaining together of these operators must not be confused with the combination of two or more persons to form a single operator(see page 106).

After having played games whose structure is the Klein group, the children will notice soon enough that if we limit ourselves to a part of the states and operators, we sometimes obtain a "sub-game" of the original game. For example, let us just take small children and operators consisting of one person only; the rule of the small boy and that of the small girl will then be all that remains. One small boy will act as the neutral element and one small girl as the "changer of direction". For example, let us say that we have two boys and three girls at the beginning, and that one little girl knocks on the table; now there will be two girls and three boys, that is, there will be one girl less than there are boys. This means that "more becomes less, and less becomes more." Multiplication by a "small girl" will thus correspond to multiplication by minus 1, whereas multiplication by a "small boy " will correspond to multiplication by plus 1.

When several persons are present in the operator the rule of each person must be observed. Fpr example, multiplication by two small boys will mean that each person leaving must be replaced by two persons of the same kind. It is the multiplication by "plus two", whereas, if there are two small girls knocking on the table, we must replace each person leaving by two persons of the opposite sex; this corresponds ~~th~~ to multiplication by "minus two". In fact we can get the answer by applying the operator corresponding to each *single* person to the *original state* in turn and then "adding the results. By means of such concrete situations the children will learn the multiplication and addition of integers as seen from within a structure that contains them. This structure, as well as the structure of the integers themselves, is known to mathematicians as a *ring*[*].

---

[*] In the formal language of algebra our ring here is the set of ordered pairs of integers with the operations of addition and multiplication defined by:

$$\text{(i)}\quad (p,q) + (s,t) = (p+s,\ q+t)$$

$$\text{(ii)}\quad (p,q) \times (s,t) = (ps+qt,\ pt+qs)\text{, different from}$$

the usual multiplications of complex numbers or of vectors.

The operators forming the Klein group under this multiplication are just the four pairs $(1,0)$, $(0,1)$, $(-1,0)$, $(0,-1)$.

It is worth noting that this ring contains divisors of zero, as may be verified from $(1,1) \times (1,-1) = (0,0)$.

Such a situation does not arise in the ring of single integers.

There is not enough space here to describe all the activities we might present to children as embodiments of abstract mathematical strctures, such as rings, fields, quotient fields, vector spaces, etc. All these structures can in fact be embodied in game-like activities, well suited to the elementary level. Such activities will prepare children for the abstract stage which will be developed at the secondary level. If the abstract and formal stage is presented without any previous introduction in terms of personal activities, mathematical structures will have no referent for the child. The child will become hostile to such learning; he will not understand how the structures he is learning are related to reality, and mathematics will always remain an alien subject to him. Consequently, it is very important that, during the elementary stage, the child should have numerous contacts with various activities which prepare him for the learning of more and more complicated mathematical structures which he will need later on. The group structures will also serve him in the study of geometry just as the study of vectors will be necessary when the study of movements in space is undertaken. Again, the study of groups also serves to formulate coherently the isometries of regular polygons and of regular polyhedra:

5.10 Logical operators

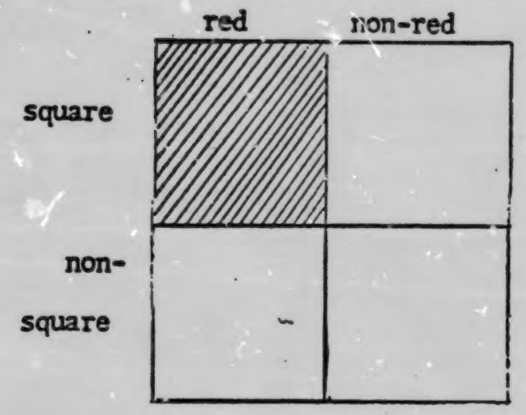
We have already mentioned the importance of certain logical operators when speaking of set-theoretical operators. In the program suggested by us, we do not present logic through the propositional calculus but through the attribute calculus.

We choose a certain universal set, for example the set of logic blocks, or the people set or the trimath , or the quadrimath which provide a certain number of variables while giving these variables a certain number of values. These variables may be shape, color, position, sex, number of holes, etc ... With such a set of objects attributes such as "red" or "square" or " large" can be embodied in a manner suited to the children's level by taking the set of all the elements of the game possessing the attribute in question. This set will be called the largest possible set associated with this attribute. After that logical operators can be used at the same time as the set-theoretical operators. For example, the negation operator can be couched in concrete terms by taking the complementary set of the set which is the realization of the attribute which we want to negate. Thus if we wish to embody "redness", we take all the red elements of the game. To negate this operator, we take all the elements of the game which are not red, then the set of non-red objects of our game corresponds to the negation of the attribute "red". The "non" used to form the expression "non-red" from the expression "red" corresponds to the operation which consists in taking, instead of the set of all the red ones, the set of those which are not red. The input of this operator is the attribute "red", the operator is the non and the output is the attribute "non-red".

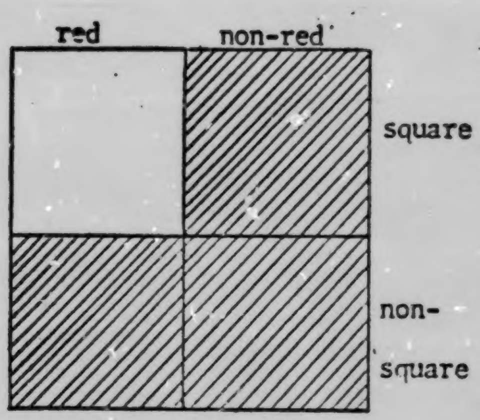
Negation is a unary operator, because only one input is needed to make the machine work.

We have already seen that the conjunction operator is expressed by the use of the word "and" . In everyday language the word "and" is omitted. For example, instead of saying square and red, we talk of an object which is a red-square. We mean that the object possesses the property of being square and that of being red at the same time. From a logical point of view it is the simultaneity of the possession of the two properties which is applied to the two properties. From a set point of view, we construct the intersection of the set of squares with the set of red ones. This means that the operator "and" corresponds to the intersection operator when we speak of sets. In the same way, the observation: "Each object is red or square", can be embodied by making a union of the set of red ones with the set of squares. The operator "or" therefore corresponds to the union operator.

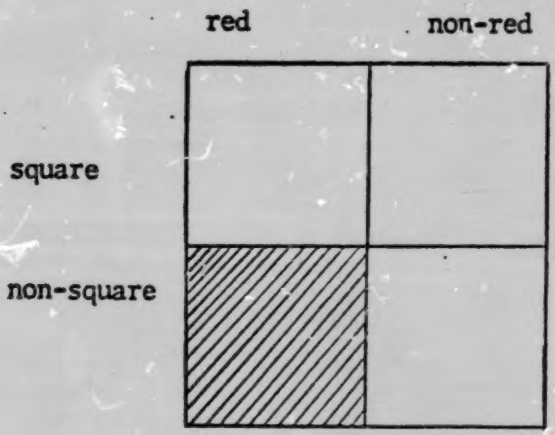
Thus we arrive at certain means of operating on logical states. These states are the attributes. We can also operate on states obtained after a previous operation and thus pile up operators on top of each other. For example, we could say that an object is not at the same time red and square. The "non" is an operation which applies to the attribute "red and square" whereas the operator "and" is the operator which provides us with the conjunction of the property of being a square, with the property of being red. The "non" operator which is applied after having applied the operator "and" must be clearly distinguished from the result of the application of the "non" operator to one or the other of the attributes followed by the application of the operation "and". This means in practice that to say that something is not at the same time square and red, is not the same as saying it is at the same time a non-square and red. Here are the diagrams showing the difference:


red
non-red
square
non-
square

at the same time square and red


red
non-red
square
non-
square

not at the same time square and red


red
non-red
square
non-square

at the same time non-square and red

The set of rules of the game used for passing from a logical state to another logical state is called attribute calculus. At a certain moment we start operating on operators and this is when we start reasoning. The use of logical operators on logical states is possible during the pre-operational logical stage; but to operate on operators and thus to reason can only be done at the operational stage of reasoning, normally called the hypothetical deductive stage according to Piaget's epistemological theory. For example, take the attribute "red and square at the same time". This attribute has been obtained from the attribute "red" on one hand and the attribute "square" on the other hand and we have used the binary logical operator "and" to obtain the attribute "red and square". We can conclude that such an object is red because all the red squares must be red. We can also conclude that they must be square; but how do we come to this conclusion? The set of all the red squares is a sub-set of the set of all the red blocks. If we know that an object is in a sub-set of a set, we can conclude that it is also in the set. Our reasoning, is based on the procedure of passing from a sub-set to a set. We "reason" in the direction sub-set ----> set. Since the set of all the red squares is a sub-set of the set of red blocks, we can establish a reasoning pattern which allows us to go from "red and square at the same time" to "red".

It is clear that when we say "red" we give less information than if we specify "red and square". We must realize that any reasoning goes from a stong statement to a weaker statement. There are much more complex ways of reasoning but most of these ways can be based on this very simple supposition: if we have an element of a sub-set, we can conclude that we have an element of the set. This principle is that of the sub-set.

We need another principle which we could call the intersection principle. As a rule we have more than one premise in a reasoning, In this case, the supposition which we make and from which we can draw conclusions, must be based on the intersection of the set which corresponds to the first supposition with the set corresponding to the second one. Suppose a person has in his hand, "either a red one or a square", then we know he has chosen a block from the union of the set of red ones with the set of squares. This can be our first premise, our first supposition. Our second supposition could be: someone has found out what our "person" has hidden and he has seen that it was not a square. Then we have two premises; we know he has chosen his block from the union of the squares with the red ones and we also know that he has chosen a non-square. Therefore his choice has been made from the intersection of these two sets because we know that each block, which satisfies these two conditions, is a non-square and at the same time either a red one or a square. We can see that this intersection contains only red ones which are not square. We can conclude that he must have a red non-square and consequently, by the previous subset-to-set reasoning, we can deduce that he has a red one, because all red non-square must be red, because the red non-squares form a sub-set of the set of red ones. Thus we have the two premises:

1. either red or square
2. non-square

We can infer, from the intersection principle, the following conclusion: red and non-square at the same time, then from the sub-set principles that is is red.

Note that all this time we have spoken only of attributes. We reason on attributes and not on propositions. In a way, we gradually introduce existential operators (quantifiers), but only implicitly, by using, each time, the largest possible set whose elements possess the attribute we wish to express in a concrete way by providing just such a set.

We must not confuse the negation of an attribute with the negation of the quantified proposition. "All the blocks are red" is a proposition. Its negation would be "it is not true that all the blocks are red"; whereas the negation of the attribute "red" would give us the negation inside the propositions, namely "all the blocks are non-red". The two propositions "it is not true that all the blocks are red" and "all the blocks are non-red" do not have the same meaning !

When learning about logical operators, these difficulties are avoided at the beginning by not making explicit the quantified propositions implied when we take the largest set possible each time we choose an attribute. The logical operators must be applied, not to these propositions, but to the attributes. The calculus of attributes differs very much from the predicate calculus. Predicates assert something about a subject. The quantifiers assert something about such an assertion. For example if we say: "The dog is bad", is bad is the predicate and the dog is the subject. But it is not a real proposition because we do not know what dog we are talking about. Therefore, we must show a dog and say: "This dog is bad", and use the demonstrative quantifier "this" frequently used in language. We can say: "Some dogs are bad"; this means we are talking of a certain universe of dogs and we quantify the assertion, "the dog is bad", by the quantifier "some". We can also say: "all dogs are bad". Therefore this also means we are thinking of a certain set of dogs and we say that in this set "the dog is bad", is a true proposition for any dog we might choose. The expressions "some", "all", "this", quantify our sentences containing subjects and predicates.

A series of cards is being prepared for children of 8 to 9 years of age, with the help of which children should learn the distinctions between subjects and predicates, quantifiers and their negations. We also refer the reader to the article "Exploitation de jeux Logico-Sémantiques par des enfants de huit ans" written with the cooperation of Béatrice Lefèvre and published in "Langue française" in December 1971.

The use of logic in mathematics and in language should be part of the essential knowledge acquired by the child when he leaves school. Curiously enough, up to now no one has bothered with the learning of logic at an early age whereas very complex computations have taken up a considerate part of the program intended for children. It is certainly very important, in order to be able to exercise accurate judgment about statements people make, to know something about the correct use of logical relations in everyday life. Everyone should know about elementary logical operators just as they should know about elementary arithmetical operators. Relationships between different quantifiers, between simple logical operators, between subjects and predicates, should be known by all so as to facilitate intelligent communication between people. The elementary presentation of these ideas should be part of any elementary school curriculum.

5.11 Geometry

We have already seen the importance of geometrical operators. General geometrical relations should not be neglected either. For example, the fundamental relations of incidence should be presented so that the child is able to appreciate in detail more complex relations between figures. The term "incidence" is a global term which contains notions such as "to pass through", "to lie on" and relations of this kind. For example, we say that a point lies on a line or that a line passes through a point or that two lines meet in a point. Here the line is linked to the point by the relation "to pass through" or "to lie on". When we say: "The point A lies on the line B", "lies on" is the relation which links the point to the line". "Line B passes through the point A" is the inverse relation. In mathematics it is said that point A is incident with line B and that line B is incident with point A. The same expression "to be incident with" is used instead of "lies on" or "passes through". Thus the incidence becomes a relation which is its own inverse, because if A is incident with B, then B is incident with A. For example if line A meets line B at one point then line B meets line A at this point.

To introduce in a concrete manner the idea of incidence, a board placed in a horizontal position on which children can place pins to make up "lines", can be used. Two pins can be placed anywhere on the board; the child can place himself at the level of the board; by shutting an eye and adjusting another pin to the other two pins, he will notice that all three pins can be well aligned until he can see only one. He can also hide other pins and while he places them so as only to see one and so that the others are hidden behind this one, his eye will be aligned or incident with the line determined by the first two pins. He can thus make other lines and he will realize that certain pins belong to more than one line. Thus, he will see that two lines cannot have more than one common point, that is, two sets of aligned pins cannot have more than one common pin. If two lines on the board, even if they were to be continued indefinitely, can never have common pins, they are said to be parallel. Thus the notion of direction can be introduced and the relation "to have the same direction" will become an equivalence relation which will lead to the introduction of parallel lines as equivalence classes. The set of all the lines in a plane, which are mutually parallel belong to the same family of lines, to the same equivalence class, produced by the equivalence relation "to have the same direction as". We will meet many other geometrical equivalence relations. For example, we might take a pin and say: "We must find other pins which are all at the same distance from this first one". In this case we will make a circle of pins. Thus we can build a system of concentric circles, the centre of the whole system being the first pin placed. These sets of pins will provide the equivalence classes produced by the equivalence relation "to be at the same distance from the first pin as".

An important relation between lines is perpendicularity. This is a symmetric relation. If a line A is perpendicular to a line B, line B is perpendicular to line A. In the case of Euclidian geometry, no line can be perpendicular to itself, which means we have a symmetric situation but it is anti-reflexive.

We recall that a relation is reflexive if for all elements the relation links this element to itself. A relation is symmetric if, each time an element is related to another by a relation, this other element is also related to the first one. We also recall that a relation is transitive if, each time an element is related to a second one, and the second one is related to a third by the relation, then the first one is always related to the third by the same relation. It can be easily verified that the relation "to be parallel to" is reflexive, symmetric and transitive and so is an equivalence relation.

"To be parallel" means "to have the same direction as".

Any given line has a definite direction and belongs to the family of lines all possessing this same direction. As no line can have two different directions, no line can belong to two different families of parallel lines. So the set of all the lines in a plane are split up into families having no elements in common. All lines in the same family have the same direction and so are related to each other by the relations

"To have the same direction as"

or

"to be parallel to"

Clearly these relations are equivalence relations.

Mathematically, an equivalence relation is often defined as one which possesses the three properties of reflexivity, symmetry and transitivity!

It is not difficult to see that such a definition comes to the same thing as our first definition, namely that of a similarity relation which splits all elements of our universe into disjoint sets of similar elements. A line always has the same direction as that same line, consequently, our relation is reflexive. Also if a line is parallel to another, this other one is parallel to the first one. The relation of parallelism is evidently symmetric. Also if a line is parallel to a second one, and the second one to a third, the first one is always parallel to the third one, because the three belong to the same family of lines since they all have the same direction and belong to the same equivalence class. Consequently, our relation is also transitive.

Parallelism is thus an equivalence relation according to the above definition, possessing these three properties. It subdivides the universe of all lines of a plane into families, the members of which are parallel to one another.

As for perpendicularity in a plane we can note another interesting property. If a line is perpendicular to a second one and if this second one is perpendicular to a third, we can see that the first line is not perpendicular to the third but that it is parallel to the third. It must be noted that this interesting relation between perpendicularity and parallelism holds only in the plane. In space it is possible for a line to be perpendicualr to a second one and a second one to a third without the first being parallel to the third; just imagine the three edges of a rectangular box meeting at one corner of the box.

The development of the study of geometrical relations which finally leads to distances and to angles and then to geometrical operators such as rotations translations and relexions, must form part of any elementary study of geometry. The study of the isometries of the plane, of ways of covering the plane, are also interesting studies which can be undertaken at the elementary level. Such studies do not present any great difficulties for the child. child.

The importance of the links between geometrical studies on the one hand and arithmetical and algebrical studies on the other hand must also be emphasized. We have already said that the study of distance is related to that of measure in general, as it is to the study of equivalence and order relations such as "as long as", "longer than", and "shorter than", etc .... The problem of measure can also be considered in a more general context, leading to the study of probability as a measure of a set of events.

In our zeal to come to grips with abstract structures, applications must not be forgotten. One of the main reasons for studying mathematics is to be able to apply it in the reality of our surroundings. It is evident that skill in arithmetic is necessary in everyday life but it is less obvious how knowledge of other structures helps us to acquit ourselves well in our modern environment. One abvious application of mathematical structures is to language, to the mother tongue as well as to other languages. Language is a complex structure and many of the relations, operators and structures discussed so far, have their analogues in language. An understanding of how language works helps us to get to grips with the business of expressing ourselves in our own language as well as in other languages we might wish to learn.

Logical competence has immediate applications in the construction of arguments or in trouble shooting, namely in finding fallacies in arguments, in separating premisses from conclusions, proof from truth and so on.

The knowledge of classification gets us used to more precise and explicit thinking, so necessary in our complex new environment. The knowledge of how to handle order relations will help us in looking for items that have been ordered according to complex criteria such as handling coding systems used in libraries, or working with a dictionary, a street directory or a telephone book. The use of double entry tables prepares us to read maps, to interpret statistical tables, to understand longitudes and latitudes, and the like.

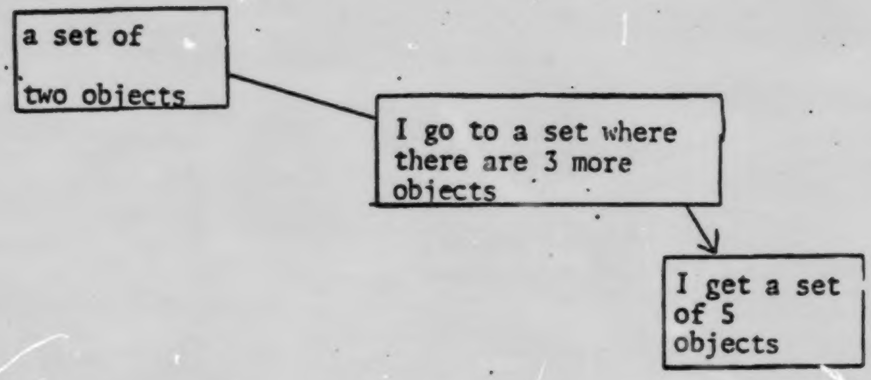
essential information from complex happenings and in comparing events with each other for similarities and differences.

The use of transformations and the expression of elements of a structure in terms of other elements, should get us used to thinking in terms of links between events, certain combinations of events determining other events. The use of a formalized language encourages us to think in terms of "bare bones" without any frills; it is often necessary to disentangle the essential from the unessential, given certain criteria.

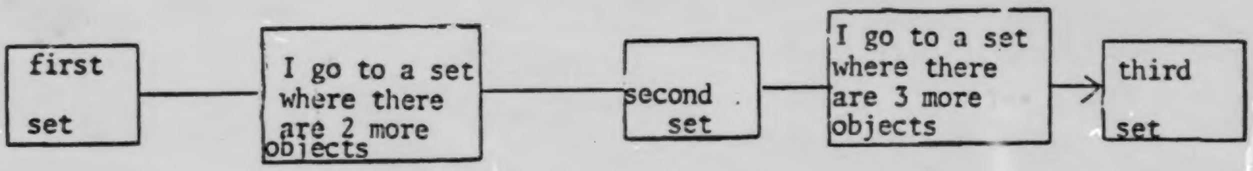
It will be appreciated that we are more likely to apply the attitudes we have developed as a result of learning mathematics, rather than apply any specific content, although naturally we use much mathematics in simple computation, measuring, estimating, making probabilistic decisions and so on.

5.12 The Links between Psychological and Mathematical Problems

We have already noted that there is a psychological barrier between the operational and the operational stage in the learning of a mathematical structure. There is a big difference between knowing that, given

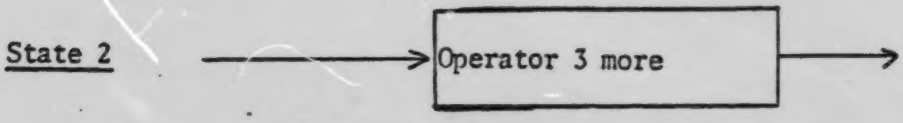

a set of two objects
I go to a set where there are 3 more objects
I get a set of 5 objects

and knowing that:


first set
I go to a set where there are 2 more objects
second set
I go to a set where there are 3 more objects
third set

in the third set there are 5 more objects than in the first!

In the first case we know that:


State 2
Operator 3 more

leads to state 5 . In the second case we know that the chain of operators

operator 2 more followed by operator 3 more

can be replaced by

operator 5 more

At the pre-operational stage of the learning of addition, the child will be able to do the first problem but will be lost in the second one. At the operational stage, he will be able to operate on the operator "2 more" by the operator "3 more" and obtain 5 more. It is very important to select the problems faced by a child as he passes on to the operational stage. The reason for the distinctions between "states" and "operators" is almost entirely of a psychological nature, rather than of a mathematical nature.

Another difficulty faced by children learning mathematics is that of situating their new knowledge in a larger context. The traditional methods did not take account of this need, the chapters in most textbooks being quite separate and independent of one another. The principle of "situating in a context" leads us to consider the order in which simple tasks will be presented in relation to complex tasks. Very often to ensure maximum transfer effects from acquiring one competence to the acquisition of another, it is better to begin with the learning of a more complex structure, of which another structure, which is the real objective of the learning, is a more particular case. It is for this reason that we suggested the Klein group or the cyclic group of four elements as an introduction to the multiplication of integers. The rule of signs for the multiplication of integers is expressed in the table of the two element group

⊗	+	−
+	+	−
−	−	+

or

	a	b
a	a	b
b	b	a

which is part of the table

	a	b	a^1	b^1
a	a	b	a^1	b^1
b	b	a	b^1	a^1
a^1	a^1	b^1	a	b
b^1	b^1	a^1	b	a

(Klein group)

and also of the table

	a	b	c	d
a	a	b	c	d
b	b	a	d	c
c	c	d	b	a
d	d	c	a	b

cyclic group
of four elements

the cycle goes as
follows

$\longrightarrow \boxed{c,\ b,\ d,\ a} \longrightarrow$

We can see that the child will thus have on hand a mathematical context, inside of which he will be able to situate his knowledge of the "multiplication of integers". The famous

minus × minus = plus

will lose much of its mystery, and will appear as one if the most natural things in the world.

We have already emphasized the importance of the stages in the learning of structures. Superimposed on the passage from the pre-operational to the operational stage, we consider stages in the psychodynamics of the detailed learning process itself, which leads the child through different levels of abstraction. These stages can be stated briefly as follows:

1. Free play. First interaction with the environment.
2. Structured play. The child realizes there are constraints, he "plays with" these constraints, he imposes them himself. He invents "games with rules".
3. Comparisons. The child compares his different games and applies discrimination in terms of the relations of inclusion, identity, exclusion and overlapping as far as his games are concerned.

In particular, he will come to understand how two games, which look quite different are in fact identical considered from the point of view of the internal structure of each game. This is a gigantic step towards thinking "in the abstract".

4. Representations. Once a child has abstracted what a certain number of different games have in common and has classified them in categories according to these similarities, he will need to be able to pinpoint each class by means of a representation.

To see at one glance what he has abstracted, he will use arrow diagrams, tree-structures, cartesian systems and the like.

5. Symbolization. The child needs a language to express the abstractions he has reached. He can develop such a language by observing the properties of his representations. This is the symbolization process.

6. Formalization. Finally, the child will have to put in order all the properties he has found in his representation. It is during this stage that he determines his axioms, his methods of "deduction", in preparation for the formulation of theorems.

The task of constructing a curriculum that will lead to the acquisition of certain mathematical skills consists in making an appropriate synthesis of the mathematical and psychological requirements: We need to reconcile the structure of what has to be learned with the needs of the person learning. Unfortunately these two aspects of the same problem are rarely considered together. The skeleton program sketched in this chapter is the result of an attempt at such a reconciliation. The mathematical hierarchies and the stages of psychological development have been considered as a whole in order to arrive at these suggestions. A more detailed account of the psychodynamics involved is given in the book "Les Six Etapes du Processus d'Apprentissage en Mathématique", by the present author and published by O.C.D.L., Paris, and in the last two chapters of "Mathematics through the senses", published by the National Foundation for Educational Research for England & Wales, London, England.

Chapter 6

New Mathematics Programs

Everybody is talking about the "new mathematics". It is not generally known that what is described as "new" is some 150 or 200 years old, having taken a Russian sputnik to get mathematical ideas past the obdurate conservatism of the curriculum kings. A few years ago it was the fashion to start mathematics projects; new catchwords and clichés were rapidly being formulated in order to regiment the unsuspecting educational public into just another straight jacket, the difference being that teachers and children now monotonously began repeating polysyllabic words such as"associative principle"instead of chanting tables in unison. Behind all the "modern" drum-beating, and the shameless commercial scramble to get in first with the shoddiest but most saleable mathematics curriculum full of the new clichés, there has here and there some serious work been done. How can the uninformed and unsuspecting public possibly tell the difference between a carefully researched, mathematically and psychologically sound curriculum and accompanying methodology, and the mathematically and psychologically half-baked commercial packages peddled round the world by profit-hungry publishers ?

It is intended to give the public just a few guidelines, not of course foolproof ones, whereby a discriminating searcher amongst the morass of new curricula and methods can be helped to make a somewhat more informed choice than he might have done without these guidelines.

Before we can suggest what these should be, some excursions into mathematical and psychological territories will be necessary. Since there are as yet no satisfactorily validated educational theories, there will be no need for any excursions into educational territories. These are largely unexplored, mainly because governments have been unwilling to finance any expeditions of any scope whatsoever to explore them.

Much research into the foundations of mathematics has been done in the past few decades, and the architecture of the foundations is becoming increasingly clear to those mathematicians who take the pains to read up the evidence. The knowledge of how fundamental mathematical concepts hang together are now beginning to have an effect on how mathematics educators think it should be taught to children. The connexion is still extremely tenuous, since mathematics continues to be regarded as a very exacting discipline, and on the whole those who concern themselves with educational problems find it difficult if not quite impossible to grapple with the simple task of apprehending mathematical truths expressed in mathematical literature. It is really the demands of industry and of the armed forces that have forced educators to rethink their educational ideas on the teaching of mathematics, and in this way the beginnings of an alliance between the two camps are being formed. Since education is not a scientific discipline, in the sense that physics, chemistry, and mathematics are scientific disciplines, the alliance is a somewhat curious one. Educationists on the whole rely on books, with a certain modicum of visual and other aids, and on a certain amount of discussion in grappling with the practical problems of introducing the new mathematical architecture into the classroom. The old, ossified centuries-old mathematics curricula which the "new mathematics" is supposed to replace are being taught according to very well defined and well established practices, known as "traditional teaching methods". These must have been developed over the decades and centuries through trial and error and thus reached some kind of practical equilibrium, producing a certain proportion of mathematicians, scientists, technicians and the like. But today these proportions are no longer adequate for manning a modern technological society. Hence all the hue and cry.

Intuitively the reformers have realized that with new curricula must come new methods. And this is where the educationalists could not help. The obvious next question is, whether the psychologists could help ? Upon looking at the abundance of "learning theories" in psychology, it soon becomes clear that very little of any of these theories is of any use to the teacher on the classroom floor. Learning theorists have been so far concerned with the kinds of learning that are extremely elementary, and so the psychological problems about how such complex behavior patterns as those that abound in the learning of mathematics can be explained became the urgent but wide open problems of a section of psychology ill equipped to meet the challenge. The net result is that now more and more psychologists are becoming interested in what they refer to as the higher cognitive processes (as opposed to "lower" ones such as the learning of strings of nonsense syllables). Contacts have been established between mathematicians, educators and psychologists (see e.g. Dienes-ISGML, 1966), but there are still too few persons available who are equally well conversant with the disciplines of mathematics and of psychology, and at the same time familiar enough with the realities of the classroom to be able to contribute realistically to the eventual solution of the curriculum and method problem.

To make matters even more difficult for the layman, mathematicians are not agreed amongst themselves as to the desirability of teaching "structure" to young children. The educators and psychologists might argue with some force that the mathematicians are not entitled to any but a lay opinion on this subject. Worse still, psychologists are not agreed about the best methods of teaching children mathematics, and as most of them do not understand very much about the overall structure of mathematics, this is not surprising. The result is the free-for-all engaged in by everyone who can profitably fish in these troubled waters. Teachers as well as parents throw up their hands and the children get more and more confused.

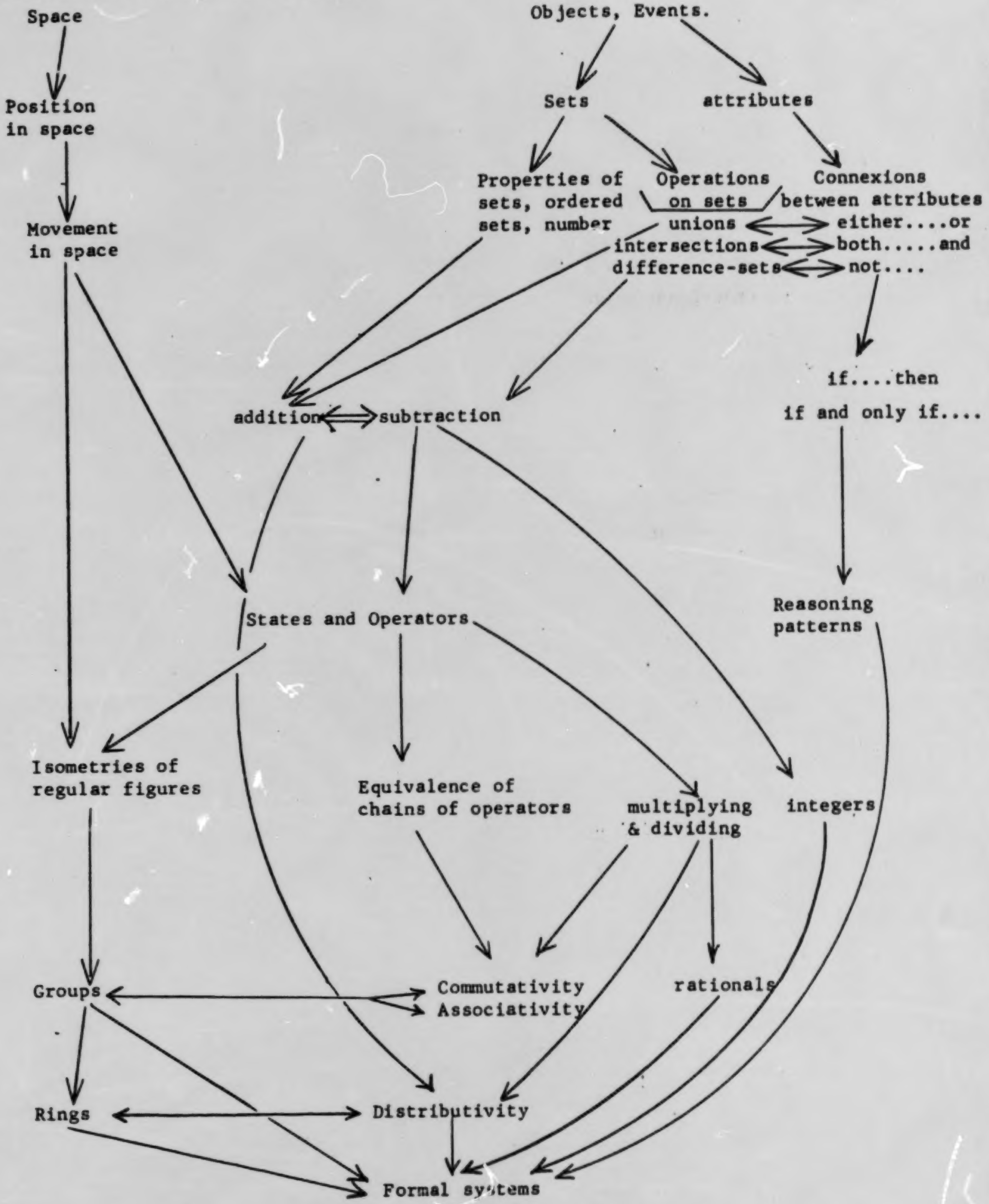
In the three school projects around which the present text centres an attempt has been made to bring psychology, education and mathematics together in a new setting. Judged against the world's school population this is a small start, but nevertheless one which has given the authors the conviction that it was worth making and could lead to a satisfactory solution of the problem of innumeracy on a much larger scale in the future.

The Mathematical Picture.

Since the late fifties, a certain number of meetings of mathematicians, educators and psychologists have taken place on the international scene as well as within the framework of national cultures. During these discussions, the argument about teaching children structure and not rote-learned responses to certain sets of stimuli, has been practically universal and unanimous (Cambridge Conference on school mathematics, 1963). It may be assumed then that the general trend of educational thinking is in this direction. Whether this will eventually prove to be a success or not, it is not possible to say. Now what are these structures? In a short note it is impossible to give an account of a subject so vast, but the following flow chart (Figure 6-1) suggests the first fundamental structures with which a child might be greeted during his first few years at school.

Figure 6-1

A flow-chart with the fundamental structures with which a child might be greeted during his first few years at school.


Space
Objects, Events.
Position in space
Sets
attributes
Properties of sets, ordered sets, number
Operations on sets
Connexions between attributes
unions
either....or
intersections
both.....and
difference-sets
not....
Movement in space
if....then
if and only if....
addition
subtraction
States and Operators
Reasoning patterns
Isometries of regular figures
Equivalence of chains of operators
multiplying & dividing
integers
Groups
Commutativity
Associativity
rationals
Rings
Distributivity
Formal systems

The main path in the structural map above takes care of the journey as far as the natural numbers are concerned. The operators of adding and subtracting on the one hand, and those of multiplying and dividing on the other, offer suitable growth-points for extending the number system to integers and fractions respectively.

Considering successions of adding and subtracting operators, it is always possible to reduce such successions to either one addition and no subtractions, or to one subtraction and no additions. Such are the considerations that lead to the extension of natural numbers to positive and negative integers. Again, we can consider successions of multiplications and divisions. Any such succession can always be reduced to a pair of such operations, one multiplication and one division. This then leads to the introduction of rational numbers or fractions. The neutral operators in these two cases are "add or subtract nothing", known as zero, and "multiply and/or divide by one" known as one. The necessity to have states on which any of these operators can operate leads to the introduction of integer states, i.e. states of "minus one", "plus three" and so on, as well as to the introduction of fractional states such as "two thirds", which for example is the name given to the state arising out of operating on the state one by a multiplier 2, followed by a divider 3 (or in the reverse order). "Plus three" is the name of the state obtained by operating on the zero state by the "add three" operator. "Minus one" is the name of the state obtained by operating on the zero state by the "subtract one" operator and so on. The combination of the two extensions give us all the positive and negative rational numbers.

In the study of movements in space there are corresponding operations, with almost identical structures. If we move ourselves in a movement which takes every point of our body along in a straight line, i.e. if we never turn, we perform a translation. Translations can be "added", since the effect of any two successive translations can be brought about by one single translation. "Subtracting" a translation could be the adding of that translation but in the opposite direction. It will be verified that all the rules of adding and subtracting positive and negative numbers are obeyed by such translations.

There are also spatial equivalents to multiplication and division. These are enlargements and shrinkings. Shine a flashlight on to a plane figure, and then let the shadow fall on to another plane which you hold parallel to the plane of our figure: An enlarged version of the figure will be seen. If the distance from the flashlight of the plane of the shadow is to the distance from the flashlight of the plane of the figure as q is to p, then for every distance of p units in the figure there will be a distance of q units in the corresponding part of the shadow. This means an operator $x \frac{q}{p}$ applied to the state of the figure will give another figure. It is possible to reconstruct the shadow figure out of some material such as cardboard, and apply to it a further transformation, say $x \frac{s}{r}$, so that for every distance r in the second figure there should be generated a corresponding distance s in the third figure. The problem of generating the third figure from the first one will yield as a result the multiplication rule for fractions.

The above are merely some examples of how the arrows on the structure-diagram can be turned into realities. Many mathematics programs, particularly as they are put into texts, seem to degenerate into a series of topics, without much attempt being made to connect the topics in any natural way.

Examples of structuring the mathematics programme with a view to emphasizing connexions could be multiplied indefinitely. It is a pity that more designers of mathematics curricula do not make such structuring one of the aims of their programme.

One other important feature of any mathematical structure might be pointed out. Apart from the connexions of a structure with other structures, discussed above, we might be interested in the generalizability of the structure, or even in ways in which a structure can be presented in a way which is as general as possible and yet within easy grasp of a child's comprehension. Much of the fertility of mathematical thought comes from the mathematicians inveterate habit of trying to generalize, or to look at things more generally. One way to help the achievement of generality is to vary all possible variables inherent in a mathematical structure. For example, a space is not just one-dimensional for a mathematician; it can have any number of dimensions. Linear operations are extended to non-linear ones, rotations about one particular axis to rotations about any axis, and symmetries about a line in a plane are extended to symmetries about a plane in space. Those responsible for either the construction or for the application of curricula could well bear such considerations in mind and see whether there are any encouragements for children to generalize as they go through the curriculum, or whether the "topics" are presented in the most general ways possible. For example, is arithmetic restricted to "tens and units", very slowly graduating to "hundreds" and so on ? How can a child generalize if the steps by which he has to make such a generalization are so few in number and so far separated in time ? And if he is given only the base of ten, how can he get the taste for generalizing ? If we close up artificially the territory he can move in, how do we eventually expect him even to want to get out of it, let alone succeed in doing so ? If we give him one or possibly two dimensions of geometry to work in, how will he ever think of an arbitrary number of dimensions (variables) ? Seeing that the generality of structures should be met with as early as possible, in order to prevent the child from making a premature closure around a particularized structure the opening of this closure, i.e. generalization, would be more difficult at a later stage. The study of number bases as a mathematical tit-bit at the age of eleven will not have the same effect as the study of the expression of number properties in any arbitrary base at the very beginning, i.e. round about the age of six.

The Psychological Picture.

Another consideration in assessing a curriculum is its treatment of the role of symbols in mathematics. For those people for whom mathematics is a meaningless empty shell of manipulative rules, mathematics can indeed be equated to the symbols by means of which it is expressed. But to those of us who regard mathematics as abstracted structure, the symbols are but the purveyors of such a system of structures. In this case we must surely be careful not to confuse the process of abstraction with the process of symbolization. If the symbol is to stand for an abstract structure, then this abstract structure must have been in some way formed or at least begun to be formed if it is to be possible for the symbol to stand for it. In assessing the way a programme of mathematics is presented it is as well to examine whether any attempt has been made to consider the processes of abstraction and of symbolization as separate though related. As far as abstraction is concerned the question whether a particular curriculum takes steps to ensure that abstraction does in fact take place is of crucial importance. Since an act of abstraction is an act of realizing the common feature of many different situations, one would expect those who construct curricula and their accompanying methodology to have at least attempted to present mathematical structures in a number of different ways.

To attempt to apply an analytical type of technique to the learning of structures that have not been properly constructed is likely to be ineffective from the point of view of both comprehension and efficiency in handling. Construction (mental as well as physical) takes place out of the component pieces, with a view to putting them together to meet with certain kinds of requirements. This is always the preparatory stage to the establishment of the insight about the internal relationships of the pieces, with possibly a few exceptions. Care should be taken that the organic nature of the concept formulation cycles is not unduly interfered with. Since different children will have different dynamics of constructing their concepts, provision should be made for individual differences. This can be done in various ways. The children may be grouped and encouraged to work from instruction cards. Alternatively, they may be put to work with a more severely structured type of programmed text. The latter, however, seems to take the life out of mathematics learning as a dynamic and exciting adventure (see e.g. Dienes 1964, Dienes-Golding 1967).

Where possible counterexamples should always be given to illustrate cases in which conditions in a theorem are not met and its conclusions can not apply. In the beginning when notions such as commutativity and associativity are being learned, it would seem sufficient to present children with some non-commutative and some non-associative situations, making these as natural as possible; this introduces them to contrast. Certain things happen sometimes and the opposite kinds of things happen at other times.

Points to look for in Mathematics Curricula.

It is now possible to give some pointers, possibly helpful to the users of the many variegated mathematics programmes and curricula, to enable them to make a less random judgement than they would otherwise have done. Our judgement, based on the considerations of the last few paragraphs, will naturally be biased. It will be biased heavily in favor of structure and against associationism, in favour of children's liberty of thought and against authoritarianism, and in many other ways which will be quite evident to the reader. If the reader should find this bias unacceptable, he is invited to compile a more satisfactory set of principles according to which a different set of suggestions could be constructed.

Here is a questionnaire in which the underlined answers are biased towards the kind of dynamic and psychologically viable mathematics programme as conceived by us. A score can be obtained for any programme by simply counting the number of answers which are the underlined ones when applied to the programme in question. The questionnaire includes questions of a general, mathematical and psychological type.

1. Does the programme mainly serve future industrial needs ?

   Yes or No

2. Do "modern words" such as set, commutativity, associativity or group mostly occur in just the one part of the programme in which they are introduced ?

   Yes or No

3. Are the instructions to the teachers so explicit as to give her no choice of action ?

Yes or No

4. Are there many references to other relevant mathematics programmes?

Yes or No

5. Are the connexions between parts of the programme stressed throughout ?

Yes or No

6. Does the programme proceed slowly through a very large number of very small steps ?

Yes or No

7. Are consistent attempts made to present structures in ways as general as possible ?

Yes or No

8. Does the learning of each structure begin with an attempt to construct, from concrete experiences at a young age and from mental ones later on ?

Yes or No

9. Are all children expected to cover each part of the programme at the same time ?

Yes or No

10. Are individual differences amongst children adequately catered for in the way the programme is organized ?

Yes or No

11. Are children presented with open ended situations as a regular feature of the programme ?

Yes or No

12. Is the process of abstraction encouraged by the presentation of the same structure in a number of different guises ?

Yes or No

13. Is there frequent use made of contrast in making clear the meanings of mathematical properties or of conditions of theorems ?

Yes or No

14. Are the ideas (i) symbolic, (ii) abstract, confused with each other in the development of the programme ?

Yes or No

15. Is provision made throughout for children to check the validity of their thinking ?

Yes or No

16. Is there an attempt to relate mathematical ideas to those in other fields such as language or music ?

Yes or No

It is suggested that a score of 10 or over is a good score, a score of 5 or under is a bad score. Again, the scoring can be altered according to the particular bias of the reader. For example a "correct" answer to question 5 might be given more than one point.

Projects also differ significantly from one another in other important ways, the rightness or wrongness of which is a far less clear cut matter. Some of these differences may be pinpointed by attempting to answer questions such as the following:

1. Which of the following does the project regard as important sources of involvement for the student ?
   a) The teacher's personality
   b) The environment
   c) Structured materials
   d) Texts and worksheets
   e) Films
   f) Sound and video tapes
   g) Group work and co-operation
   h) Competition

2. The nature of mathematics calls for some hierarchical ordering of its presentation. Does the project treat this ordering as;
   a) A closed sequence of stepwise exposition, as in programmed learning ?
   b) A sequence of prompts for investigations which lead the student to discover principles for himself and gradually put them into a logical order ?

3. Any group of mathematics students, however homogeneously composed, will sooner or later reveal a distribution of ability.
   Does the project
   a) Adjust its approaches to meet this distribution ?
   b) Provide a minimal sequence, closed or otherwise, which the weakest students are able to complete ?

4. Does the project
   a) Believe in keeping the students of a class together by periodic evaluations ?
   b) Allow the individual student to assimilate sequences at his own rate and take an evaluation when he is ready ?

5. Does the project
   - a) Use the environment to promote the students' general appreciation of the value of mathematics ?
   - b) Use the environment to further the abstraction of mathematical structures ?
   - c) Use the environment for the application of mathematical structures which are already known ?

6. Does the project use modern topics
   - a) In isolation from traditional topics ?
   - b) In conjunction with traditional topics wherever this is appropriate ?
   - c) For their psychological value in helping to develop powers of abstraction, of generalization and of their mathematical mental process ?

7. Does the project in the teaching of a given topic
   - a) Set a deliberate limit to the choice of possible approaches ?
   - b) Liberally borrow from what other projects have to offer ?
   - c) Try to select approaches to suit the individual student ?

Our two questionnaires are designed to provide the reader with a set of ready made criteria for assessing the three projects which are about to be described and for comparing them with one or more of the projects that have tried to solve or reduce the learning problems of children in the elementary school.

An attempt to classify the modern mathematics programmes.*

A series of world wide efforts to gather relevant information on the part of the International Study Group for Mathematics Learning - three meetings of mathematicians, psychologists and educators, one at Stanford in December 1964, another in Paris in April 1965, another in Hamburg in January 1966 - culminated in the publication of a study by the UNESCO Institute for Education, Hamburg (Dienes - ISGML, 1966). In this the distinguishing features common to the new programm are classified into five categories:

1) Emphasis on structure,
2) Emphasis on meaning, understanding and discovery,
3) The creation of a positive attitude towards or even emotional involvement in mathematical activity as a basis for motivating mathematical learning,
4) The content of the curriculum to be broadened far beyond the conventional four rules,
5) Induction to precede deduction wherever possible.

Against this common background the study goes on to distinguish among the projects six different types, depending on whatever features appear to be most strongly emphasized. In its own words the classification is "largely arbitrary", but for the purpose of this book it adds pertinent detail of content and approach to the last section. The classification is also rapidly becoming out of date. Most projects now pay at least lip service to sets. Few may now be called arithmetically oriented . However, it may still be worth while to examine the present situation in at least similar terms.

---

* See Appendix 6 - 1 at the end of this Chapter for a list of the more prominent modern programmes.

The sixfold classification of approaches runs as follows:

1) The set-orientated approach,
2) The arithmetically-orientated approach,
3) The geometrically-orientated approach,
4) The science-orientated approach,
5) The symbol-game-orientated approach,
6) The multiple-embodiment approach.

The set-orientated approach

Over the last hundred years mathematicians have become more and more aware of the fact that the concept of a set is fundamental to the understanding and development of every branch of their subject. To match this, learning research has produced evidence that the young child has a rudimentary notion of the set concept before he understands number. The set-orientated approach which adapts the curriculum to these new standpoints is found, among many others, in the Greater Cleveland Math Project, the School Mathematics Study Group, also in various programs that have grown out of the Ball State Math Project. Naturally it has received a large number of interpretations and treatments, some of which have been criticized for doing little else besides stating the obvious in more cumbersome language.

The arithmetically-orientated approach

This takes many of the relations between numbers and sets for granted and starts children with the (set of) natural numbers as being intuitively known. It works within the framework of a more traditional curriculum and concentrates on making the operations of arithmetic more meaningful and acceptable by employing Stern, Cuisenaire, Unifix or the Multibase Arithmetic Blocks as structural apparatus.

No examples are quoted in the UNESCO Report. Because of the growing tendency to widen the field of mathematical experience for young children beyond arithmetic it would seem fair to regard this approach as obsolescent.

The geometrically-orientated approach

A good example of this approach is to be found in the Minnesota Mathematics and Science Teaching Project. The isomorphism between real numbers and points on a line is used for introducing children to the natural numbers. An easy visual representation of addition and subtraction is thereby obtained and extended to multiplication and division later on.

An approach relying heavily on geometry has also been used in Belgium by Professor Papy. Children at the elementary levels can be brought into contact with rigorous proofs without being troubled by the symbolic language in which non-geometrical proofs would otherwise have to be presented.

Finally it should be noted that in the spirit of Felix Klein's Erlangen Program geometry appears in many modern programs as transformation geometry. The Euclidean property of congruence is then simply that property of a figure which remains unchanged by translation, rotation and reflection.

The science-orientated approach

This approach exploits the close connections between science and mathematics in the teaching of both subjects, graphs and measurements being obvious examples. It figures prominently in the Minnesota Mathematics and Science Teaching Project, also in the California Elementary Science Project. In Europe one of its chief exponents is M. Naisky at the Lycée Mangeron near Paris. Naisky attempts to simulate the properties of numbers on special machinery constructed from a meccano kit.

The symbol-game-orientated approach

Here it is taken for granted that the learner has passed from the manipulation of structural materials to the stage where he is able to think mathematically with the help of symbols alone. The approach occurs extensively in projects such as the University of Illinois Committee for School Mathematics, the Madison Project and the University of Illinois Arithmetic (later Watertown) Project.

A simple type of game that can be played with open sentences (equations) is the following:

$$\square + \square = \square \times \square$$

Substituting the numbers 3, 5, 2 and 4 appropriately will give a true statement. Who is quickest at finding another ordered quadruplet? Again, symbols can be made to represent machines with an input and output. One possible objective for a game is then to ask for combinations of machines which can produce a given output from a given input.

The multiple-embodiment approach

This is central to all the projects discussed at length in separate chapters of this book. Put crudely it is an extension of the well-known fact that if you want a child to learn the meaning of the word "red" in a foreign language a good approach will be to show him several red things which differ from each other in everything except their colour. The procedure singles out what is essential and has been used successfully in helping children to abstract mathematical relationship correctly.

If the approach receives its strongest emphasis in projects based on the theories and research findings of Dr. Dienes it is also implicit in any programme that stresses variety of explanation for a given topic or induces the student to solve the same problem by more than one route.

Traditional versus modern approaches.

Our evaluations[*] were designed to contrast a "traditional" with a "modern" approach in mathematics teaching. As we have seen, there is no such thing as the modern approach any more than there is such a thing as the traditional approach. There are two groups of educational methods (with materials) which place a greater emphasis on different educational objectives (e.g. problem solving versus gain in knowledge), but the differences are often difficult to notice, for instance if a "modern" orientated" traditional method is compared with a "traditionally" orientated or conservative modern method. We could only expect to find significant differences in the outcomes of the students' performance if there were also "significant" differences between the educational environments of their schools.

Throughout the remainder of this chapter we shall describe in outline the Projects involved in our evaluations. We shall also point out those environmental variables which - according to our judgment - could be responsible for the expected differences in the students' cognitive and affective states.

The Budapest Study.

In the Budapest Study this variable could well be an authoritarian versus a democratic teaching method, the latter laying more emphasis on learning through discovery and problem solving. Of course "traditional" instruction does not necessarily consist of an authoritarian taskmaster issuing streams of words unintelligible to his apprehensive students sitting in rigid order and in silence, repeating on demand formulas they have memorised without understanding (Crittenden 1969, p.35). On the other hand, learning through discovery and problem solving, when used as an exclusive method in modern mathematics projects also have several disadvantages in that they provide less systematic knowledge, make an uneconomic use of time, and so on.*

---

[*] For details see Volume II throughout.

There is a variety of learning methods which can be used in every learning situation. Although we favour the discovery method, we can agree with Crittenden (1969,p.31) that "there are some kinds of learning in which the discovery (and problem solving) method is either clearly less satisfactory than some others, or simply does not apply. For learning how to do a whole range of things - reading, writing, swimming, cooking - in which physical and practical skills are involved, it seems that the most effective procedure is a combination of instruction, demonstration, practice, and imitation. This is clearly so,if our objective is to learn such skills well and with a reasonable economy in time." However, as far as the learning of mathematics is concerned, we do not feel its main objectives to be the acquisition of certain skills.

The Sherbrooke Study.

In the Sherbrooke Study we noticed another objective, additional to those already mentioned. This was the effort to integrate the regularities often found in mathematics with the regularities that can be found in other areas of knowledge.

Until recently not very much effort has been made in this direction, and a few serious attempts have met with severe criticism. But we can well understand the emotional basis in the Report of the Provincial Committee on Aims and Objectives of Education in the Schools of Ontario (1968, p. 69) when they ask the question:

---

* For a systematic approach to what is learned in mathematical discovery, see e.g. Roughead and Scandura 1968.

> "Should all pupils be taught by means of logically organized and separate courses in traditional subjects such as reading, spelling, arithmetic, history, science, literature and grammer, or should all pupils enjoy the stimulation of lively ideas and be given ample opportunity to discuss them, with the satisfaction of learning by discovery? "

Perhaps we are unable to find a suitable method of how to connect these different subjects, possibly because we have not as yet discovered the way in which to arrange the educational environment that enables all the children to enjoy the process of information acquisition. Maybe we have not discovered the situations which could make it possible for the pupils to discover the "secrets" of our world. But it certainly does not mean that we do not have to try to do all these things. In his book Mathematics through the Senses Dienes (1973a) has tried to encourage the beginnings of co-operation between disciplines as dissimilar as mathematics, physical education, dancing, music, and art. As can be seen from this book (p.1) and even more in the everyday life of the schools were this method is followed, it is possible to create situations in which the children can learn some mathematics through moving their body, through touch, through their sense of rhythm, through their appreciation of color and shape; they can personally enjoy the elusive process of abstraction by going through its various stages from the most concrete and simple games to the purest of mathematical abstractions (ibid. p.2).

The New York Study.

The most important difference between the New York Study and the other two was that the two schools involved in the comparison were both following a modern mathematics program of the Dienes type.* In both shhools the program was well established, with the result that the variable to which different performances and attitudes could be attributed did not arise from the internal environment of the schools but from the socio-economic background of the pupils who attended them.

The school with which the New York school was being compared was a school of French-speaking Canadians in Sherbrooke, Quebec, where Dr. Dienes had his headquarters at the time. The Ney York school was a private, fee-paying school and the Canadian school a public elementary school. There were thus considerable differences in the professional, cultural and economic status of the parents, differences which, according to many authorities, can profoundly affect the progress of children in direct as well as indirect ways.

---

* The Study assumed this form when the control school in New York running a traditional mathematics program unexpectedly withdrew from the investigation.

APPENDIX 6 - 1

Some important projects teaching modern mathematics.*

AMERICA
The United States.

Program	Year Started	Present Range	Essentials
University of Illinois. Committee on School Mathematics.	1951	Secondary Grades 7, 8	Logical structure of mathematics; study of patterns; consistency; precise terminology. Learning through discovery. Early verbalization discouraged. Non-graded units to be used in accordance with student background and experience.
Madison Project	1957	Secondary Elementary	Discovery of structure through finding one's own solutions to problems. Discussion through careful questions. Unstructured tasks; social application deliberately omitted. Ungraded material intended as a supplementary program. Integration of arithmetic, algebra, and geometry.
School Mathematics Study Group	1958	Secondary Elementary	Structure of mathematics. Concepts of mathematics as part of the whole of mathematics. Spiral approach. New content as well as conventional topics. Opportunity for conventional practice and review. No strong position on verbal-operational question.

---

* Based mainly on Lockard, 1972. For some more useful information see e.g. Rosenbloom, 1964; Lockard, 1972; Sherman, 1972; Ashlock and Herman, 1970; Lamon, 1972.

The United States.

Program	Year Started	Present Range	Essentials
Arithmetic Project	1958	Elementary	Mathematical exploration. Topics with "travel" - an adventure. Intuitive thinking: guessing; inventing; trying things out. Developing a feeling for mathematical ideas. No stress on verbalization; not a complete course of study; no grade levels. Project seeks teachable alternatives for important mathematical ideas.
University of Maryland. Mathematics Project.	1959	Grades 7, 8 Elementary	Mathematics as a language; precise terminology. Properties of a mathematical structure. Unifying concepts; particular emphasis on number systems. Learning through discovery. Inductive and deductive reasoning. Verbal and operational components simultaneously encouraged. Fundamental learning processes a major concern of the project.
Standord University. Sets and Numbers Project	1959	Elementary	Concept of set and operations on sets. Precise language;mathematical laws. Logic. Relationship between set theory and foundations of arithmetic. Algebraic and geometric principles.
Greater Cleveland Mathematics Program	1959	Secondary Elementary	Logical structure of mathematics. Mathematical laws. Discovery approach. Search for patterns and relationships. Emphasis on a continuous and systematic flow of mathematical concept formation. Grades K - 12.

The United States

Program	Year Started	Present Range	Essentials
Minnesota Mathematics and Science Teaching Project	(1958) 1961	Elementary	Three specific mathematics structures: the real number system, Euclidean space, a space with measure. Concepts of an algebraic structure and a deductive science. Connections between mathematics and science; cultural and historical aspects of mathematics. Working toward a coordinated mathematics-science curriculum K-9 and undergraduate courses for pre-service education of teachers. Ungraded units. Grew out of the MINNEMATH program which started in 1958. Project director 1961-1966.
Committee on the Undergraduate Programme in Mathematics (CUPM)	1959	All college mathematics	Study & improvement of college curricula.
Comprehensive School Mathematics Program (CSMP)	1965	Grades K-12	Games, tapes, films and other media. Vertical acceleration through independent study. Stochastic determination of goals.
Fleming School Project (New York)	1969	Grades K-6	See chapter 10

CANADA

Program	Year Started	Present Range	Essentials
Centre de Recherches en Psycho-Mathématique	1966	6-14 years	See chapters 8 and 9

CANARY ISLANDS

Canary Islands Mathematics Project (CIMP)	1965	5-16 years	All possible structured materials in use Experimentation, logic & creativity superseding drill

AUSTRALIA

New Guinea Primary Project (Dienes)	1963	5-11 years	Adaptation of Sherbrooke type program to local New Guinea conditions.

EUROPE

Belgium.

Centre Belge de Pedagogie de la Mathématique	about 1960	Elementary Secondary University	Development of mathematics materials for all levels

HUNGARY

Program	Year Started	Present Range	Essentials
OPI Mathematics Project (Hungary)	1963	Grades K-9	See chapter 7

ROMANIA

The Systematic Teaching of Probability Dr. Efraim Fischbein Institute of Psychology Str. Frumosă 26 Bucarest Romania	1968	12-15 years	To add the topic of probability to the curriculum and teach it from a psychological point of view

THE UNITED KINGDOM

Midlands Mathematical Experiment (MME)	1960	-----------------	Insight into structure Preparation of students for uses of mathematics in industry and science
School Mathematics Project (SMP)	1961	11-18 years	To update the secondary curriculum and maintain liaison with practising mathematicians

THE UNITED KINGDOM

Program	Year Started	Present Range	Essentials
Nuffield Mathematics Teaching Project	1964	5-13 years	A wide-ranging contemporary approach Usually materials are given to teachers who develop them for the children
Bulmershe Project	1970	5-11 years	Adaptation of Sherbrooke type programme to English conditions.

Chapter 7

The OPI Mathematics Project in Hungary

Origins and Evolution

After about two years of preparation (Varga 1964), in September 1963 a school mathematics project was started in two classes of a school in Budapest, Hungary. Initiated by Tamas Varga, the project has since been developed into a nationwide scheme, the OPI Mathematics Project, spreading at present over one hundred classes in a dozen Hungarian cities. The OPI Project basically set itself the same task - but with much more limited opportunity - as the Madison Project, that of "Seeking the best experience with mathematics which can be provided for children at the pre-college level" (Davis 1964).* Its start is well described by Varga himself:

---

*Though the OPI Project developed mainly from the concepts of Dienes, it has also assimilated certain ideas from educationists working in the same direction all over the world: Sawyer in the USA, Davidov and Elkonyin in the Soviet Union, Krygowska in Poland, and others. See also Suranyi et al. 1972.

The evaluation of the Project (Budapest Study) was undertaken with the support of the Orszagos Pedagogiai Intezet (OPI: Hungarian Pedagogical Institute) and the Pedagogical Department of the Technical University of Budapest. A short description of the methods and results was prepared for the World Conference on Education (Asilomar, Calfornia 1970) and was also delivered at Columbia University Teachers' College (May 11, 1970) and at the Institut National de Recherches et de Documentation Pédagogiques in Paris (28 May 1972). For the Hungarian form of this short version see Klein 1972a.

First beginnings

"Pilot work in the teaching of mathematics first started in Budapest in the autumn of 1961. After a two weeks' summer course, many teachers expressed their agreement and sympathy to the new approach. But when it came to the question of doing some teaching based on the same principles, each of them had some excuse. They clearly found the new ideas dissimilar to anything they had experienced before. Thus, we had no teacher at our disposal. We did have two sets of Multi-base Arithmetic Blocks, a fair amount of ideas about what to do, and the eye of many people on us (not all of them kindly disposed to the announced experimental teaching). Should we retire? I decided to begin myself with a class of 36 seven-year-old girls -- an age group totally outside my earlier teaching experiences. The authorities were far from being enthusiastic about the project. Eventually I was permitted to teach extracurricular material (especially non-decimal arithmetic based on the M.A.B.), but it was made clear that the official syllabus had to be covered as well. The class teacher, a self-conscious lady particularly proud of her methods of teaching arithmetic, was ordered to be present at my lessons without ever interfering. There she sat the entire year, and tried not to smile at my blunderings with children who were more favorably disposed than she. They liked playing games instead of waiting their turn to answer a question posed by the teacher. They liked the freedom to leave their places (in order to change material or worksheets). They definitely preferred this to sitting immovably on their seats. They also liked changing their ideas during the games without being punished for breach of discipline. This freedom was permitted them only three times a week, each time for a period of about eighty minutes. They availed themselves of the opportunity eagerly. The class teacher deemed that the freedom was excessive and that this jeopardized her educational efforts during the rest of the lessons. I could not doubt the veracity of her statement; very different educational practices being used in close juxtaposition are always likely to lead to conflicts.

The result of this first trial, after the first year of teaching, was the refusal of the authorities to give permission for another. The children, and a number of parents, were sorry. The class teacher took notice of the decision with a certain relief. To be frank, so did I."

A New Start

"In the 1962-63 school year, the M.A.B. material served mainly as a plaything for adults. During this time, however, the climate regarding non-standard teaching experiments in mathematics underwent a considerable change. This was mainly due to the International Symposium on Mathematics Teaching held in Budapest (27 August - 8 September, 1962). It had to be realized that the usual slogan "The teaching experiences of our best teachers should be diffused and generalized" was no longer sufficient in our age of rapid changes and that radically new departures had to be made. I was invited to start similar work again, this time in two first grades (6 plus year olds) of another school. My duty was to design the syllabus and the teaching materials, also to instruct and supervise the teachers.

This was a great opportunity and I felt I had to avoid taking too much risk. I decided therefore not to impose anything upon the teachers. They had to be convinced of the potential usefulness of what they were going to do and never to jump into darkness. Both teachers had much good will and they were well qualified. Neither of them clung rigidly to her views concerning either the syllabus or the teaching methods but they would not have given up willingly their feeling of security developed during their years of school teaching. They were instructed by word of mouth and by various written materials, including relevant chapters from Z.P. Dienes' Building Up Mathematics and L.G.W. Sealey's Creative Use of Mathematics in Junior Schools. Various efforts were directed towards enabling them to visit Leicestershire -- where much the same sort of learning as we were trying to establish had been initiated in many schools. They have not been successful as yet, though it is likely to be realized next year. The opportunity to visit these schools was, however, granted to me. I spent one month studying the work of the Leicestershire Mathematics Project by the courtesy of the local educational authorities."

The Syllabus

The traditional first grade syllabus in Hungarian schools covers numeration (only decimal, of course) up to one hundred, addition and subtraction up to twenty, with simple problems on the different aspects of these operations, the use of the " = " symbol in the sense "gives as result" and some weights and measures. As explained earlier, I had to be cautious in not departing from the usual syllabus too much; this is especially important for the first year. Striving for greater generality was the main aspect of the alterations:

1. Children were led through appropriate exercises to the grasp of " = " as a symbol meaning "the same number as".

2. "Equal" was a border case for them between "greater" and "less". They used the symbols " > " and " < " before they got acquainted with the symbol " = ".

3. By the use of Multibase Arithmetic Blocks they were led towards the general idea of place value notation rather than just to the decimal notation as a special case.

4. Through the use of Cuisenaire rods, their attention was directed toward some aspects of addition and subtraction independent of symbolism and of the particular numbers involved.

The work with these two classes was to be continued for at least three more years, and more venturesome digressions from the official syllabus were planned for these years. On the whole, without having broken away from the teaching tradition of their own and of their fellows, our two teachers gradually became aware that methods quite different from the accustomed ones are conceivable and, perhaps, even preferable. Consciously or not, they also began to appreciate some similarity between their relation to the children and mine to them, both of which can be characterised by Polya's words: "Suggest it, do not force it down their throats".

Further development

The organisers of the OPI Mathematics Project emphasized their opposition to a rapid, unfounded development. Four years later, Varga (1968) continued to write:

> "Gradual extension is a key word for us. We are against conclusions of the kind that " The results of our experiments show that our way of teaching mathematics is better than the usual, therefore it should be introduced into mass education by the school-year n/(n+1)!" By what reason and right would we impose the reform on teachers if we are against imposing mathematics on children? What we would like is:
> 1. to acquaint as many teachers as possible, as deeply as possible, with our way of teaching mathematics, or rather with some version of it, and
> 2. to attain that every teacher be given the opportunity (including the right and material aid) to practice an up-to-date teaching of up-to-date mathematics."

In spite of this, the quantitative development of the Project was impressive (Fig. 7 - 1) . In 1979 permission was given by the Ministry of Education to introduce its methods into every Hungarian school in which conditions were satisfactory.

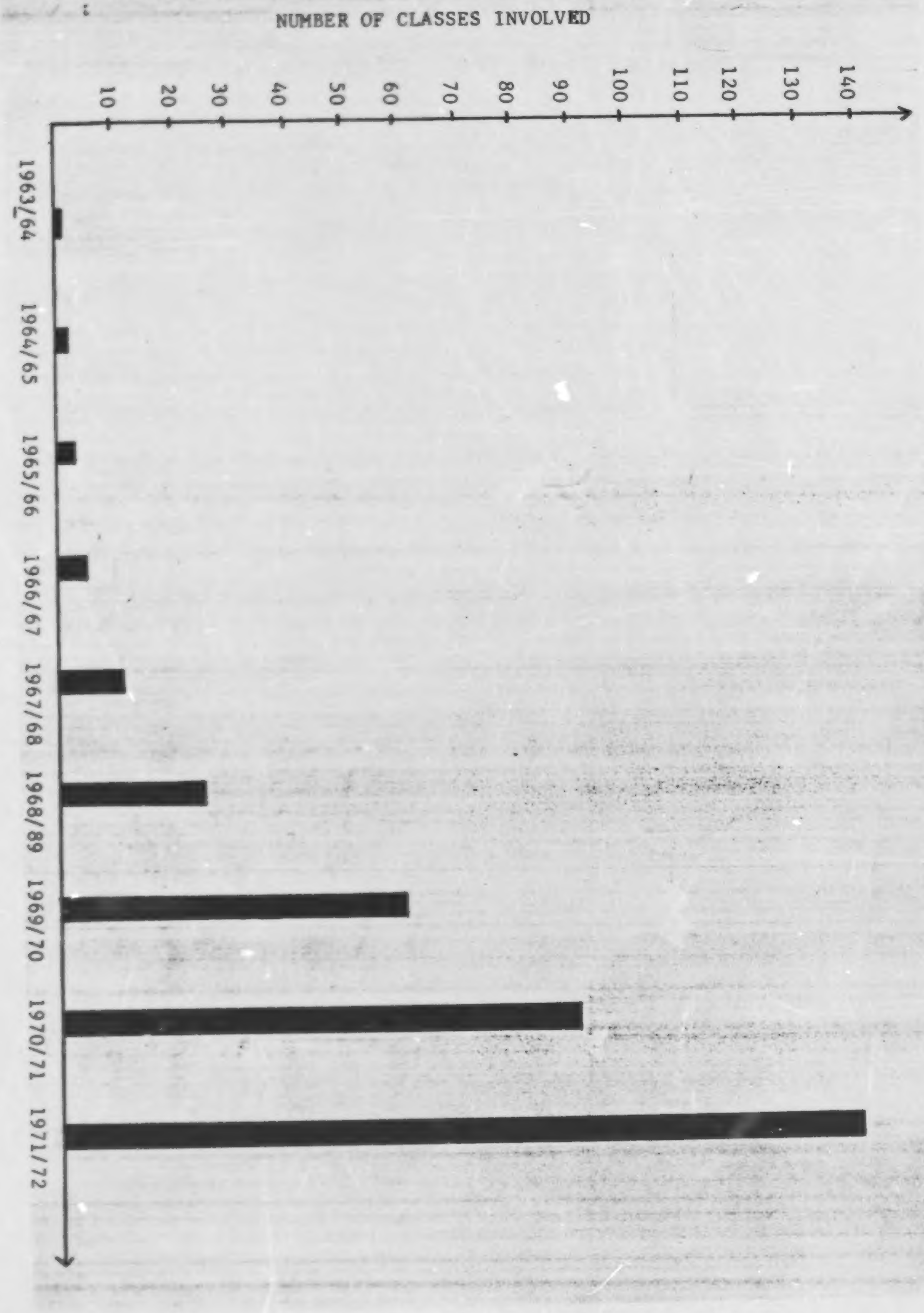


Fig. 7 - 1. The Growth of the OPI Mathematics Project.

Principles and objectives

The philosophy adopted by the OPI Mathematics Project may be summed up as "A break with the usual fragmented approach, and a view of mathematics as a whole."

Every item in the traditional curriculum is queried and a basically new set-up suggested. Ideas traditionally belonging to algebra and delayed until Grade 8 are suggested and largely adopted as early as in Grade 1,and developed according to children's abilities and needs. Many key ideas of mathematics, some of which are quite unusual in the primary (or even secondary) grades, appear in the OPI method in the first few years.

Algebra and Geometry are just as well represented in the curriculum as elementary ideas of topology, information theory, set theory, logic, probability and statistics, all at the level of children, six, seven, eight, or nine years of age.

The criteria on which the selection of the subject matter is based are the following:

1) The subject matter must be fundamentally important from the point of view of the whole of mathematics. For example: equations, inequalities, sets, functions with tables or functions from graphs, functions with formulae are all fundamentally important mathematical ideas.

2) The subject matter should be easily accessible to children through problems and through play. If an entry to the subject matter could be arranged through the use of the children's tactile sense, we have always considered this an advantage although not as an absolute condition.

3) The subject areas must be in intimate relation with each other and they must reinforce each other and thus help to construct an organically developing system.

4) The subject areas to be introduced are to prepare and to introduce the teaching of other important subject areas.

5) The subject areas must have use in modern applications of mathematics. For example, for the understanding of the functioning of electronic computers, it is absolutely necessary to be familiar with systems of numeration different from that based on 10.

Varga continues:

> "It goes without saying that we believe in the so-called spiral curriculum. Every subject matter, which might have been brought down from high school or from the university to our first and second grade children, is regarded as a first stage leading towards a more rigorous treatment later. It must also be realized that we include the ordinary mathematical activities such as the learning of multiplication tables. As a matter of fact, our children seem not so much to learn multiplication tables as to discover them. So they find the process more pleasant. We have not placed a great deal of emphasis on the improvement of fluency in giving answers to multiplication facts. We have placed a great emphasis on the extent of children's ability to work things out, to reason and to discover relationships. We feel that after such preparation, the further pieces of mathematics can be based on a much more solid foundation. In the traditional elementary school only the arithmetical operations together with

a few questions on mensuration and some verbal examples are required. This is a gross underestimation of children's abilities at the elementary level. It is clear to all who have studied the problem that a great deal of change will have to be made in the elementary mathematics curriculum. It is becoming more and more certain that such changes, at any rate at the elementary level, will be impossible without considerable and fundamental changes in methodology. The main issue is to make mathematics sufficiently challenging for children, so that it is worthy of a thinking effort. Sooner or later this pays dividends. If there is no thinking effort, the memory is likely to be over-burdened. We might lighten the burden of the student by giving him more thinking to do, as it will save him from remembering too much. But to act in this way we must change both the curriculum and the method.

Nobody need fear for our experimental guinea pigs, nor worry lest we drain the last ounce of energy from them in order to squeeze out some experimental results. They get practically no homework. They feel that during the lesson they are merely playing. In a sense, of course, they are right, although if one looked at the curriculum,one might be tempted to feel that the children were terribly overburdened. The curriculum includes such topics as arithmetical operations in the binary, ternary, four base, five base, and six base systems which are performed with ease by our young guinea pigs. It is often considered that such topics are difficult even at the secondary stage. To make matters worse, they use very long numerals, consisting of five, six or even more digits. They simplify fractions, they add, subtract, multiply and divide them, they solve equations with one or two unknowns, they work with inequalities, they examine tables of values for functions, they complete such tables, they learn representations of points in a plane through rectangular Cartesian coordinate systems and, in addition to all this shocking amount of work, the dreaded topic of set theory is introduced. But anyone who looked inside the classroom would immediately realize that no child there was anywhere near a mental breakdown."

No pressure on teachers or children

Teachers joining the scheme are given material and suggestions that may and do enable some of them to break radically with the traditional curriculum and old fashioned approaches. However, those who feel uneasy with too sudden a change are not urged to switch over at once. The OPI Mathematics Project advocates what Robert B. Davis of the Madison Project has called "light touch policy" in dealing with children.

This is especially important at the start when the teachers are teaching a great deal of material with which they have had no contact whatsoever before. " For example, the teachers have had to learn how to handle arithmetical operations in different bases of numeration. They had no previous instruction to prepare them in the experiment. The preparation has taken place on the classroom floor and the problems have been faced as they have arisen. Both the curriculum and the methodology were new to them. We adopted, for the start, a permissive attitude towards the teachers. Any new activity about which they felt uncertain they were advised not to engage in until they were certain. This permissive attitude has resulted in eventual full cooperation from the teachers who have taken an active part not only in the execution but even in the planning of the program. The positive attitudes towards mathematics which children developed very quickly as a result of the new methods went a long way to persuade teachers of the efficiency of the new methods " (Varga, 1965).

Abstraction the interiorization of concrete experience

The implications of Jean Piaget's developmental psychology for mathematical education are taken into consideration essentially as interpreted by Z.P. Dienes (1959; 1960). Varga (1965) writes:

> " In our experiment perhaps the primary educational principle which we have been using is that the process of abstraction is an essential one for which considerable time must be allowed. Children do not abstract automatically.
>
> We believe that it is only possible to abstract from the concrete, and in order to abstract, the child will have to get familiar with a great number of situations. Mathematics is very abstract. This is in fact its greatest strength, since it means, in effect, that it condenses into itself the essence of a great number of concrete phenomena. In order to get at this very abstract structure, we have to start at a very concrete starting point and lead children towards abstraction by making sure that a sufficient number of varied concrete experiences are provided for them.
>
> At the initial stages very concrete experience means sensori-motor experience. Tactile experiences are most important. In this connection our principle of abstraction implies that operations with symbols must be preceded by operations with objects. But, of course, there must be a two-way traffic between objects and symbols. It is also necessary for children to know how to interpret symbolic work, applying it to the object world. Although initially the activity is from object to symbol, later the

> activity from symbols to objects must also be encouraged.
>
> This is, in a way, what Piaget means by interiorization. All this means, in practice, is that we are putting tools in the children's hands. A distinction must be made, however, between tools used by the children and tools used by the teachers for demonstration purposes. Demonstration tools are for showing children certain relationships. The use of a working tool by each child insures that each child, as far as possible, will discover these relationships for himself."

Independence training

It is suggested that children should be allowed to make their own choice of learning activity and depth of problem solving, that they should be encouraged to construct their own problems, and control their own interaction with other children. A variety of ways of organizing the class contributes to this end. Individual work, or work in small groups with concrete materials and/or worksheets leads them to independent discoveries; seminar-like discussion led by the teacher with the whole class or a part of it can either initiate new ideas or contribute to consolidating and systematizing what has been achieved so far.

Independence training is an important tool in the OPI Mathematics Project. As Varga (1965) writes:

> "No one would quarrel with this principle. On the other hand, not everybody would interpret it in the same way, or would make use of it in the same way. Our children appear to know when they have learned something and therefore they know when they are ready for the next task. All the participants feel that there is much hidden potential yet to be tapped. Recent work with children below average ability level convinced us that the type of expansion of mathematics teaching which we have accomplished is applicable in a wide field."

2.8 The importance of individual differences

It is widely known that in a normal class, children of the same chronological age can be at very different "mental ages". Another familiar idea is that one and the same child can be inferior in terms of one ability and yet be superior in another ability. The pity is - as Taylor (1971) pointed out at the XVII International Congress of Applied Psychology - that our schools appreciate less than one-tenth of the approximately 96 different abilities known today. The Project has tried to cover as many abilities that might bear on mathematics as possible, and has let those children who are able to do more in a field produce more in both quantity and quality. We hope that, if we find a wide enough variety of methods, every child can be a high achiever in at least one mathematical field or kind of activity.

The problem of the individual differences is of course, intimately connected with the principle of independence training. With more freedom, individual differences are automatically better looked after. Those children that are able to do more will produce more, in quantity and in quality. Those that produce less are also satisfied since they also have solved problems as well as it was possible for them, having regard to their knowledge and ability.

We do not mean, however, that by educating children to be independent, we automatically look after their individual differences. We merely mean that conditions for solving the problem of individual differences have been improved. For example, we can get to know children much better, observing them working individually or in small groups than by talking to them from the teacher's desk at the head of the class. In organizing this more independent kind of work the teacher has a different task from that of the conventional teacher. He needs to provide the quicker ones as well as the weaker ones with work appropriate to their ability. It must be said that there are still a great number of problems we need to solve in this area of our experience, one of the main problems being classroom organization.

We have already said that children work individually or in small groups. But we work with the whole class, as in the case of games and team competitions. Traditional classroom activities such as mental arithmetic also have their place. Children also invent their own verbal problems, which will fit certain numerical operations that might have been written on the chalkboard. Children can work individually with their working tools as we have already seen, or they can work through the use of workcards on which they can draw or write. Such work can be preceded by verbal discussion about the contents of the workcards, but in the ideal case, there will be no need for this. The instruction cards speak instead of the teacher.

Internal as opposed to external motivation

This last sentence leads directly to the next principle. The Project wishes to find situations in which every child can find the place where he will enjoy working (or playing - because in this sense it is very hard to discriminate between the two phenomena). As Tamas Varga (1965) expresses it: "Every child is curious. He wants to seek out the secrets of the world in which he lives. Every child likes to learn. This is his nature, just as he likes to eat while he is healthy. While a child receives suitable food, he will remain healthy and will retain a healthy attitude towards eating. Many parents ruin their children's eating habits by being overanxious. In the same way many teachers ruin their children's learning habits by being overanxious."

Fortunately for our physical health, eating problems are less frequent than learning problems. In school we succeed in making learning so unpleasant as to put most children off it altogether, sometimes for life. We regard this to be the rule and we assume that a child does not want to learn so that teachers must, by rewards or promises or by force, induce children to learn and acquire the knowledge which is deemed necessary for him to have. A great deal of energy is consumed in this conflict between teachers and child; in fact the conflict has become institutionalized, and it has been termed "keeping discipline." It is kept in being by the present system of extrinsic motivation, based on grading or marking.

So prevalent is this extrinsic method of inducing children to learn that the majority of teachers would not believe that any other method is possible. A hundred years ago many school masters would have found it difficult to believe that there could exist schools without corporal punishment. This is now admittedly possible and thought even desirable by the majority of people in certain countries. It has taken a hundred years or so to make this step.

Similarly, the process of replacing extrinsic by intrinsic motivation is going to be slow. It would be foolish to discontinue extrinsic motivation at once without any transition, but it is possible to begin a slow shift of emphasis from extrinsic to intrinsic motivation. We aim at a situation in which the pupil considers more important what he learns than what his teachers or his peers thing about him for learning. This means quite simply that we try, as far as possible, to build on the pupil's interest and not on his self-concept. On the other hand, all this does not mean a devaluation of such virtues as duty and discipline. To come back to the comparison with eating: To eat with a good appetite is not contrary to regular and disciplined eating. In fact, the two go together. Likewise discipline and regularity in learning are not contradictory to learning for the love of it.

One consequence of the application of the principle of intrinsic motivation is that it becomes undesirable to allow the weaker pupils to feel that they are stupid and likewise undesirable for the better ones to feel that they are very bright. The emphasis ought not to be on their own persons but on their mathematics, on the job in hand. Children can get used to checking and correcting their own work. Something is correct not because the teacher has said so but because it has been checked and found correct. Of course the teacher should still look at the work and say whether it is correct or not. But there should never be any moral edge to such a judgement. The greatest intrinsic reward both for teacher and child is when both get on happily with the topic in hand. Nobody should be ashamed of having committed a mathematical error. To have committed an error gives an opportunity for discussion and should never be used by a teacher for making a child look small.

Of course whether a learning process is internally or externally motivated depends not only on the tools which are involved but probably much more on how they are actually used. The mechanical use of rods and blocks as simple calculating machines - which is what they may well become in the hands of unimaginative or unwilling teachers - would certainly require the use of external motivation. As Biggs (1962) writes: "If the child does not sense that there is a problem 'there', and that it is a matter of cognitive urgency that he go ahead and solve it, there is little possibility of the flash of insight occurring, and the process would thus appear to be externally rather than internally motivated. On the other hand, where the child does 'see' the connection between symbolic and concrete structures, we are probably justified in assuming the process to be internally motivated (p. 161)."

The OPI Mathematics project tries to help the teacher create an environment for the children in which they can alternate between concrete and abstract work according to their needs, thus remaining in contact with the problem and yet maintaining a high level of interest in it.

It is probably quite clear to the reader that the principles of the OPI Mathematics Project are very interconnected. For example, internal motivation is impossible without taking individual differences into account, winning the teachers over to the cause, using the right concrete material for the abstraction process, and so on. This is the reason why the project has its own "face" despite its eclecticism.

The Taxonomy of the OPI

Recently those responsible for the Project made a serious effort to overcome the many difficulties involved. By way of a start they summarised their aims in the cognitive field in a more structured way (See Table 7-1). We hope that this important first step will be followed by a corresponding summary of aims in the affective field, written by the leaders of the Project. This could then form a firm foundation for the development of a system of behavioral objectives.

TABLE 7 - 1

The Taxonomy of the OPI Mathematics Project in the Cognitive Field.*

	ASSESSING. DEFENDING OR REFUTING IF	CONSTRUCTING
PRODUCTIVE	1. A statement makes sense 2. A statement is true 3. A problem is clearly defined and data are sufficient 4. A problem contains redundant or contradictory data 5. A symbol, definition or a suggested way of solution is suitable 6. A piece of reasoning is correct 7. A solution satisfies given conditions 8. A solution is reasonable or practical	1. Formulating a problem 2. Guessing 3. Planning a solution, finding the tools 4. Finding objects, concrete or mathematical, satisfying given conditions 5. Finding all objects satisfying given conditions 6. Forming a definition 7. Developing a proof 8. Generalizing, extending by analogy, etc.
	UNDERSTANDING	KNOWING, DOING, USING
RECEPTIVE - REPRODUCTIVE	1. Seeing and distinguishing relationships 2. Understanding statements 3. Following reasonings 4. Re-wording and translating between words, symbols, graphs, etc.	1. Terminology, symbols 2. Mathematical statements 3. Routine problems, algorithms 4. Manual skills
	ANALYSIS	SYNTHESIS

* personal communication from T. Varga

2.11 Curriculum*

The organization of the pupil's work seems to be the sphere most urgently calling for radical changes. Yet it would be wrong to be satisfied with this and to regard the questions of curriculum as of minor importance. Curricula and methodological problems cannot be separated. The decisive role of the new experiments in the establishment of a uniform picture of mathematics has been pointed out. To display this picture is the primary aim of more recent endeavours. It has also been pointed out that even the most up-to-date curriculum implemented by inadequate methods may lead to a fatal formalism. The effect of this interrelationship may predominate in many other aspects as well. The traditional curriculum is always attached strongly to traditional teaching techniques and teaching practice.

Since the problem of curriculum construction has been the subject of discussion all over the world we attempt to outline here some of the key points, paying special attention to those which are not generally treated in traditional curricula and teaching; we show why their omission amounts to a serious misplacement of emphasis. Most of the topics in our list will be found in any up-to-date curriculum, and some of the more advanced concepts have been introduced into experimental curricula and pilot teaching.

---

* Suranyi et al. 1972.

Fundamental to any secondary curriculum are the real numbers. To give just one illustration: The elements of mathematical analysis can only be built on an understanding of their properties, yet these properties often remain concealed from the pupils' understanding. The gradual development of clear number concepts must thus become an urgent priority. Besides natural numbers, negative numbers and fractions may appear quite early on, together with operations, inequalities and the like. Rational numbers must form a fully clarified sphere of knowledge for the child in the primary school. Real numbers are concealed in measurement, especially in approximations. If the problems relating to them are not allowed to remain hidden, and the pupils are taught to find their solutions with genuine understanding, then it can be hoped that the concept of a real number will have been adequately clarified for children aged 14-16 years. Relations and functions also start in Grade 1, the latter in the sense of mappings, including geometrical transformations. Geometry today stands in an increasingly close relationship to other branches of mathematics. One has only to think of the linear vector fields. The general problem of the introduction of measure springs from geometrical problems and its complete solution in geometry also is given by the methods of analysis. The problems of continuity form an important meeting point for geometry and analysis. A new branch of science has sprung from the investigation of problems of this nature, namely topology. Elementary topology is a source of many surprising results, and hence also of motivation for mathematics. The integration of geometry into mathematics as a whole by no means contradicts the emphasis of its intrinsic values. The consequent development of space-perception for instance, should be continued. Combinatorics and the theory of *probability* appear at present rather late in the curriculum and little importance is attributed to them. As we see it, combinatorial exercises should have a natural place as early as in the lower grades of the elementary school. The development of the ideas pertaining to probability can also be started very early. The time available and other demands determine where and how much of it may be included in the curriculum. It must be pointed out that a distorted picture of the significance and power of probability calculus is given if its use is restricted to a small number of exercises to be solved by the application of combinatorics.

The role of mathematics in establishing *logical thinking* has been often mentioned. The curricula, however, never go beyond the recognition of this role. It is necessary to point out that the task is to make the pupils conscious of the logical structure of mathematics. This does not necessarily imply the teaching of mathematical logic as one of the branches of mathematics. Whether mathematical logic ought to be taught at all is a matter to be decided upon later, on the basis of the consideration of the given possibilities. One of the most effective principles for organising mathematics into a unity is provided by the study of *sets*. This must be gradually developed from the earliest stages and in such a way that the *organising power* of sets and their relations can be exploited throughout all other branches and stages of the subject. Another principle of organization is offered by the concept of *algebraic structures*. It is quite obvious that the introduction of certain algebraic structures and their more detailed study will be included in any up-to-date curriculum. At present it would be hard to say how and to what extent they will be included. In more recent applications, mathematics resting on completely elementary principles has played an important role. Most of this could be taught in schools but would take us far beyond the scope of today's curriculum. To decide what and how much should be taken from these application into a new curriculum will be a task for the future. It must be emphasized that the most important point is to give the pupils a clear appreciation of the basic ideas and to provide them with a fully understood knowledge that can be easily increased. In keeping with these demands the pupils ought not to be taught more than they are able to absorb.

This outline can be regarded as the first attempt to categorize the problems. Even if it is acceptable today it may require alteration after a short time. However, it will by then have served as a starting point to solve everyday problems. Accepting tentative amendments as though they were final or rejecting them altogether because of their shortcomings would be destructive of our fundamental expectation, that more viable and therefore more lasting modern curricula will be developed through a long succession of trials, errors and improvements.

2.12 Teaching aids (facts and desires)*

The situations to be analysed mathematically must, especially in the beginning, be set up in concrete materials, which pupils can see, touch and re-arrange. A large variety of such materials is now available on the educational market; many of them are flexible enough in their design to lend themselves to use in a large variety of mathematical situations. Some of these materials have become classics through their versatility and simplicity, for example, the coloured rods of Cuisenaire, the multibase arithmetic blocks and the logic or attribute blocks of Dienes, the nailboard (Fig. 7 - 2) or geoboard of Gattegno. These can and must be complemented by others if through richer experiences we wish to ensure that the child is readily led to abstraction. It might also be useful if the teacher designed materials of his own to promote his teaching, possibly with the pupils themselves playing a role in this work.

---

* Suranyi et al. 1972.

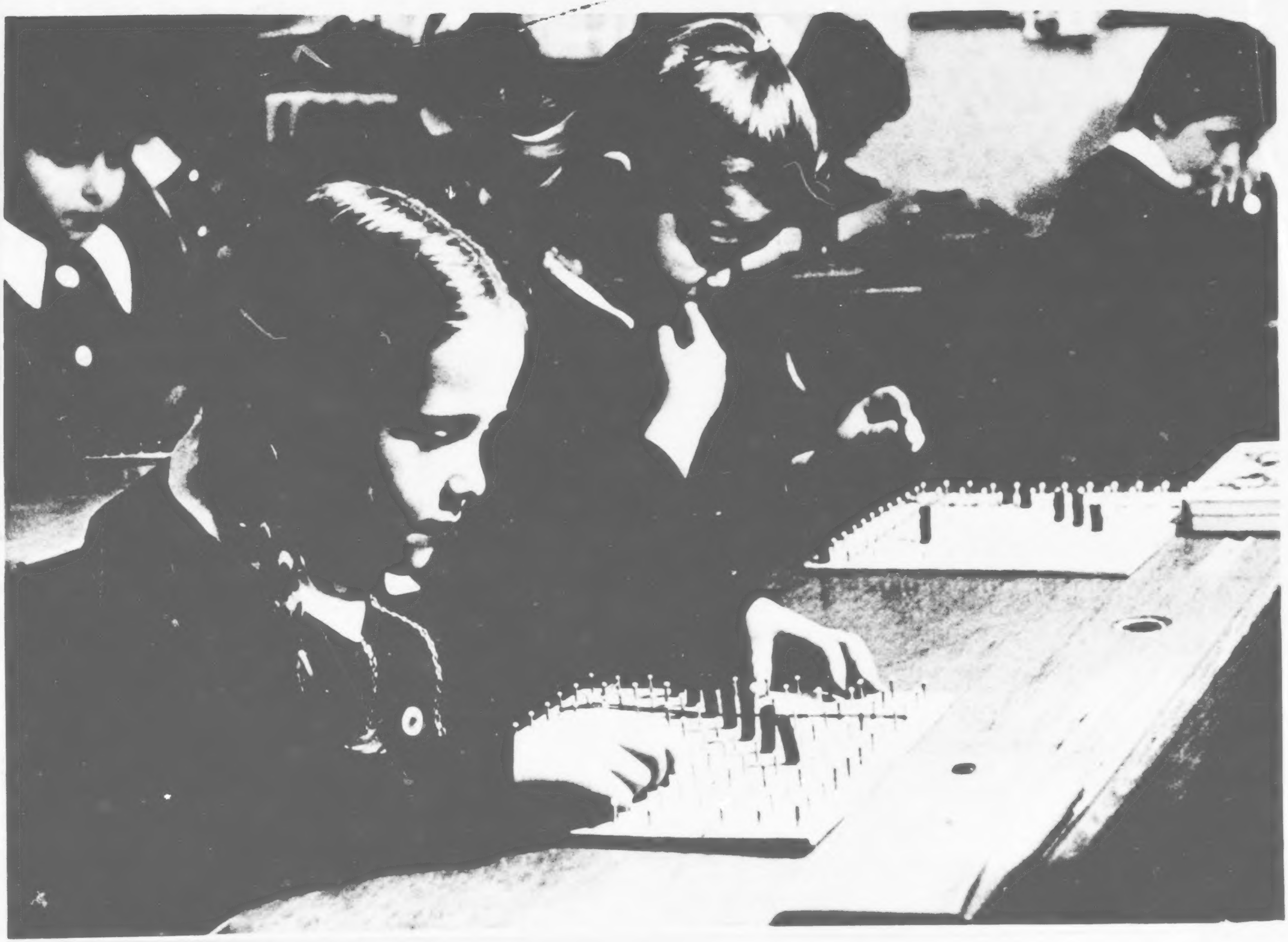


Fig. 7 - 2. Working with the nailboard.

Mathematical experience centering around manipulative activity is of primary importance, especially at the lower levels. Drawings and written symbols become more important as the new knowledge is related to the mathematical concepts already learned, to numbers instead of to quantity, to geometrical figures instead of to objects, and so on. But pupils should remain aware of the rich variety of concrete meaning that relates to mathematical sentences, whether uttered in words or expressed by formulas.

At the level where pupils work with symbols and drawings worksheets have proved very useful. These can be of great variety and serve many purposes. It is advisable to have them available in large numbers, and never to let them become a rigid system of chores. Teachers can design new ones themselves, specially adapted to certain emerging problems, to certain shortcomings, or for helping certain pupils. For the rapid production of such worksheets it is necessary, of course, to have available the necessary labour and machines.

Although experimentation seems to have concentrated on the lower age ranges it seems reasonable to assume that an appropriate supply of aids will promote independent work at any level of learning. If aids assist motivation, understanding, assimilation and retention then they ought to be of use wherever learning is in progress, though of course the particular forms they assume must change as the child grows to maturity.

The problem of the more classical aids of teaching guides such as the textbook, still remains unsolved. At present, the textbook is mainly used by the teacher to provide a supply of exercises. Although more recent textbooks and teachers' manuals have improved the situation, the disregard of textbooks is still a little worrying. Ways of exploring written information must be probed because it is in book form that the bulk of knowledge is passed on from one generation to the next.

It may be that the structure and style of the textbooks are responsible for their disregard. We have to endeavour to use a style intelligible to our pupils. The textbooks should not be confined to presenting mathematics in its final perfection. They should rather throw light on how ideas evolve; they should give the pupils a lead, explain the most important ideas and their power and encourage the more advanced reader to analyse texts on higher levels of abstraction, precision and conciseness.

With all that granted we still have to reckon with some inevitable intrinsic difficulties. The mathematical textbook is eo ipso more concise than the majority of literary works or books dealing with humanistic disciplines. One must learn even to read literary works in order to become more receptive to what the work has got to say apart from its actual plot. This is much more true for the reading of mathematical books. One of the initial steps can be to use the textbook in the course of the lesson: The pupils can read or listen to one piece and then discuss what it was all about, put questions and finally give an account of the content in their own words.

Prominent also among the new methods and aids is programmed learning and the use of teaching machines. The use of programmed books has spread both in school teaching and in individual learning. The majority of the above mentioned worksheets can be regarded as programmed aids; at least it can be maintained that they bear a close relationship to them. The applicability, realization and limitations of these methods are again subjects for much further investigation.

Even prior to such research, some points must be considered. Programmed materials take many forms. The tendency of imparting knowledge to the reader mostly through correct answers is in direct opposition to the previously outlined idea of learning through mistakes. The experience gained from the recognition of the errors plays an enormous role in mathematics learning. The application of the multiple-choice system in the teaching of mathematics is limited,among other things, by the fact that it discloses the nature of the answer to be sought, the identification of which very often plays an important part in the solution. Another drawback is the possibility of selecting the correct answer without actually solving the exercise. The danger of the programmed materials lies in their tendency to set the pupils on a road leading to one given sequence of thought, to give them only one possible way for solution, at the same time limiting their independence and their initiative. Nevertheless, they can lighten the burdens of routine and repetition for the teacher and provide an orderly introduction to a particular part of the subject which the pupil may have missed through personal absence.

The role played by variety has been emphasized many times in this book. The complex task of teaching can only be performed through the application of numerous other methods (Fig. 7 - 3). Accepting any single method or approach in the hope that it will be the final solution

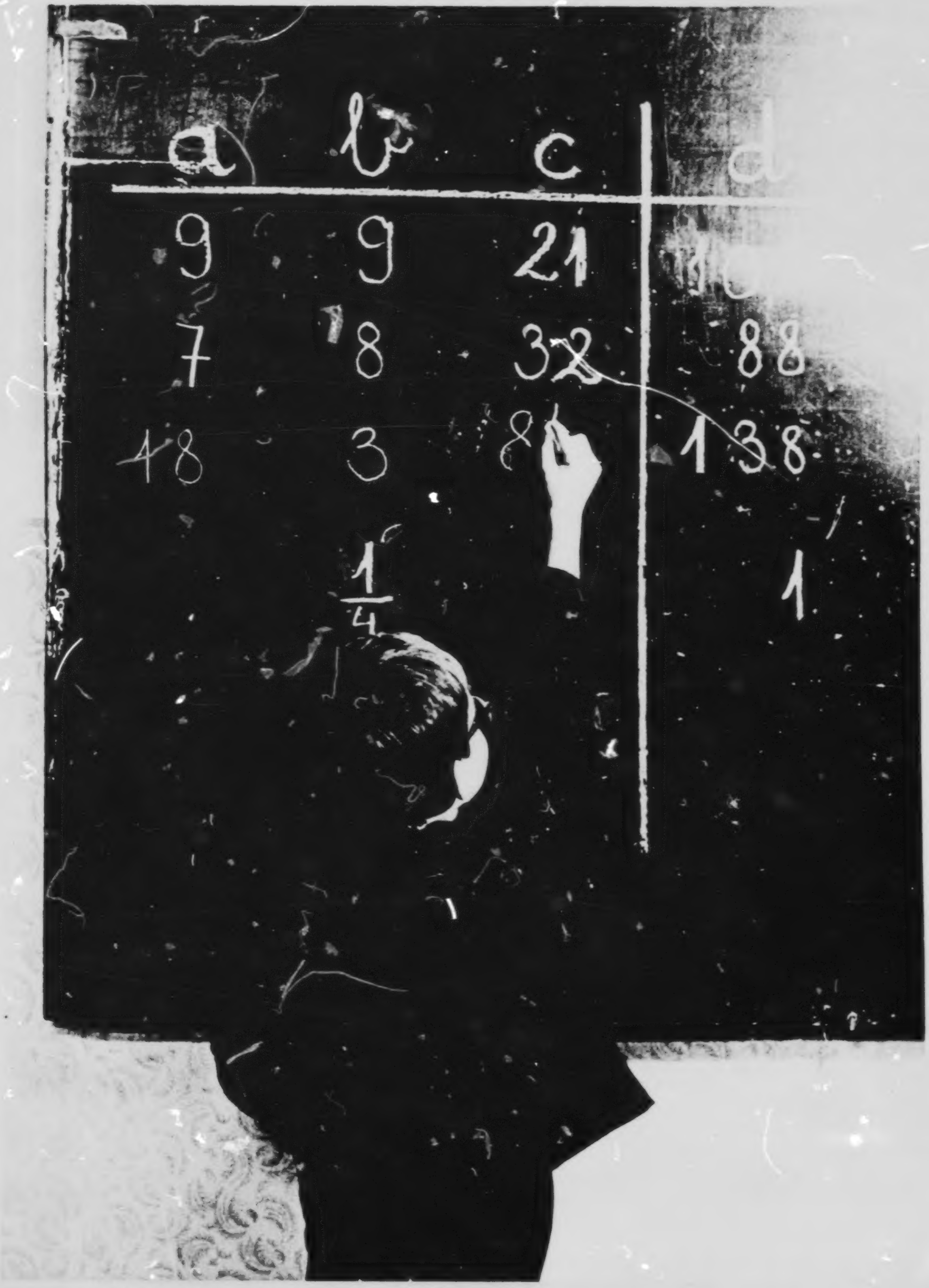

a
b
c
d
9
9
21
7
8
32
88
48
3
8
138
1/4
1

Fig. 7 - 3. The "traditional" blackboard can be used for modern mathematics as well.

of the problems of teaching mathematics is certainly harmful, especially if the approach is believed to be in line with modern ideas and up-to-date technology. This may be the case with much that is sold as programmed learning. It may become popular through momentary results mainly because it does not require too much fresh knowledge from the person who chooses to apply it. But it can never ensure a harmonious and multilateral development of mathematical skills and knowledge, and despite its partial results it may sidetrack both the teacher and the learner from seeing mathematical truths as they really are.

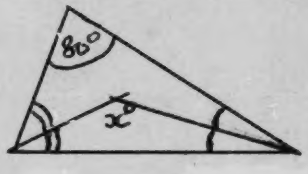
The use of tables and other devices such as the slide rule should also be mentioned in this connection, although they do not directly provide mathematical information. The majority of the pupils will need to be able to use such devices later on. Their regular use in class is therefore important, to say nothing of the fact that they motivate the pupil by enabling to achieve reasonable numerical accuracy at high speed. Where conditions permit it should also be possible to introduce older pupils to the power and speed of some electronic methods.

2.13 A Geometry lesson in an OPI Mathematics Project School

A vitally important step in evaluating an educational project is understanding the environment it creates to reach its objectives. There is no room here to describe fully this side of the investigation; the environmental characteristics we considered important are found in Table 3.

To get a feeling of the climate, we include here a description of a lesson given at the very beginning of this project.*

The lesson begins with the teacher sketching the figure below on the board, doing so without saying a word. During this time a pupil is working on a problem at the other end of the board; the problem has been selected by him from amongst others written out on cards.



After a minute or so several pupils begin to put their hands up. The teacher waits for a few more hands to go up, after which he calls on one of the pupils.

"In the big triangle one of the angles is 80°, so the sum of the other two must be 100°", declares the pupil, "The sum of their halves is 50°, which leaves 130° for the angle x."

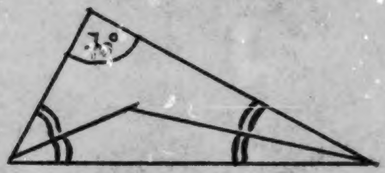
The teacher sketches a similar problem, this time with 70°. He says:

---

* Written by T. Varga (1963)

"Do this one by yourselves and write down the answer."

He strolls around the pupils' desks, and notices the following figure in one of the note-books, which he proceeds to draw on the board.



T. "What do you think of this figure?"

P. "It isn't right. We don't know whether they are all equal."

T. "If they were equal, how large would each one be?"

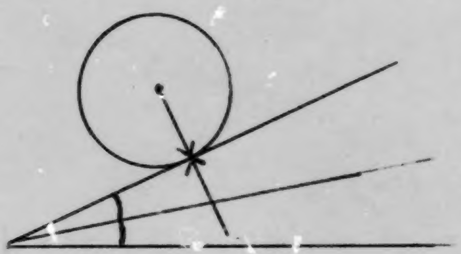
There is some general discussion, two different values are proposed; the teacher does not take sides; it is clear that the pupils are quite used to this; they do not expect the teacher to adjudicate on points at issue as some kind of oracle; they are not even looking at him to obtain a hint from his expression. The discussion in which the teacher acts only as a kind of catalyst is soon over and a decision is reached. In the meantime even the slower ones have solved the problem.

A little later again the teacher turns to a child working at the blackboard:

T. "Are you ready? Well, let's hear it."

On the board the figure drawn below can be seen, roughly sketched by the pupil who had been working there.

P. "I had to find a construction for a circle which touches a certain circle at one of its points, and also touches a certain straight line. Two conditions needed to be satisfied. One condition is satisfied by those circles whose centers lie on this line. (Here he points at the straight line which passes through the given point on the circle as well as through the center of the circle.)



"The second condition is not enough in itself, since a lot of circles could touch this straight line. So I added a further condition that the circle to be constructed also had to touch this straight line." (Here he points at the straight line touching the circle at the given point). "The set of centers of those circles which touch this line as well as the given one is the bisector of this angle." (Here he points at the bisector of the angle on the figure).

At this point somebody puts his hand up, but the teacher asks him to wait for a while. The pupil continues:

P. "There is one point of intersection of these two sets of points, which satisfies both conditions." (He points at the two sets of points while he is speaking.)

T. "Any remarks?" (A large number of pupils wish to be heard.)

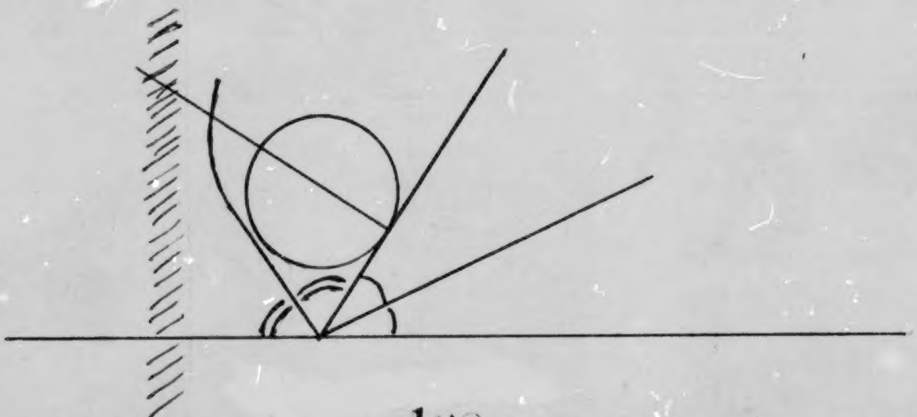
T. "Claire, what have you to say?"

Claire: "He never finished the construction. He should have drawn a circle whose radius is the distance between the point of intersection and the given point on the circle. The point of intersection is the center of the circle."

T. "Any further remarks?" Several pupils are very excited at this point and demand to be heard. The teacher calls on one of them.

P. "The problem has two solutions. These two straight lines (she runs to the board and points) make more than one angle and there is another bisector. These two bisectors together form the second set of points. So the first set and this set have another point of intersection.

She goes and completes the diagram on the board. But the second intersection would already lie off the board. She is worried.



The teacher comes to her rescue and says:

T. "Let's cheat a little bit", and draws the figure (shown on previous page). "Any further remarks?"

P. "Now they are not outside each other but inside..... the given circle will be inside the circle to be constructed."

Chapter 8

DIALOGUE AT SHERBROOKE

A good way of describing something novel is through the use of the question and answer technique. In this chapter, Professor Sandor Klein from the Technical University of Budapest poses questions to Dr. Dienes, who answers as best he can. The dialogue ranges over the historical background of the work now in progress in Sherbrooke, as well as over many theoretical and practical problems connected with the design of efficient and creative mathematical learning situations. These embody the principles and suggestions for curriculum construction as outlined in Chapter 5.

Klein: The first question I wish to ask you is if you could describe to me a little how this whole Centre began here in Sherbrooke. Maybe we could go back a little further in time and ask you what made you think of the idea that mathematics education needed to be changed, and how you did set about getting reform under way.

Dienes: Well, the whole thing is a very long story and it goes right back to the very beginning of my career when I was interested in the foundations of mathematics. In fact, my doctoral thesis was written on the foundations of mathematics, and what is more, from the point of view of an intuitionist approach. This means that mathematical ideas should be based on the intuition of the natural number. As Kronecker has said, "God made the natural numbers and man-made the rest." On the other hand, Borel and his followers made up what they called the "realist school"; that is, mathematics originates from reality, and as apperception of the real world and the extraction therefrom of certain regularities is what makes mathematics, in contrast to having the Platonic view that ideas have always been there somewhere, are eternal and just need discovering, so to speak. According to Borel, you make them up in order to suit yourself. Even according to Kronecker, all that is given by God is the whole numbers and the rest you make up yourself. So this is the kind of basis on which I tried to explain some of the foundations of mathematics.

This is one side of the story. The other side of the story is that I was always aware of how people missed the really rather beautiful architecture of mathematical relationships. How things hang together is such an essentially human experience, yet it has been essentially impossible for most people to appreciate this. I could never believe in the existence of a special mathematical ability in spite of all the psychological data; ability to do mathematics could all well be due to the way in which the subject is approached, which is of course entirely unsuited to most people. Anybody who knows mathematics realizes how stupidly it is taught. From time to time, I have come across people, particularly children, who have had difficulties and I have tried to help them. The kind of realist attitude which I picked up from Borel gave me the idea of making a reality out of mathematics so that children could appreciate such a reality and then abstract from it afterwards.

I was often asked to help private pupils who had difficulties in school; what I did with them was not problems that were in the book, but games which I had made up so that they could understand what was in the book later on. It was only after two or three lessons that the children began to realize that they weren't just playing games, that there was something there to do with the book.

In the middle 1950's, I became acquainted with some of the people in Leicester, trying to do something in the schools in mathematical education. One of these was Len Sealey. He showed me around the schools and I saw lots of children playing lots of games. They Had train tracks and they were measuring the time taken for an electric train to get from here to there. They were using trundle wheels for measuring distances. They were using buckets of sand and all this kind of thing. I thought, "Well, this is fine.", but I began to wonder how in a completely unstructured situation children would ever get beyond a certain first order set of concepts, such as, wolume, speed, time, and how to measure them. In mathematics, abstractions are piled upon abstractions and most of mathematics has to be learned from other mathematics. This seemed to me to be lacking, and so I had a little "think" about this and, in my inspection of the school situation I found that there were two yawning gaps in the understanding of children.

One was their understanding of arithmetic in general which I tracked down to their lack of understanding of place value. They didn't realize that for example in 327 the 3 really means the second power of 10 taken three times. Maybe they could repeat it after you like parrots when the inspector came, that those were really three hundreds and the others were two tens, but they didn't really know in their bones what that meant, and not knowing this, they had difficulties in learning arithmetic. The other difficulty was their complete lack of understanding of algebraic manipulations. For instance most children appeared to have no idea of parentheses. What is (2 + 3) x 7 as opposed to 2 + 3 x 7? They did not seem to know. These two sets of relationships had to be the targets for the first attack on the problem. So I designed the multibased arithmetic blocks to cope with the first one and the algebra experience materials to cope with the second one.

There was absolutely no thought in my mind to construct a mathematics programme or to build any psychological theory of learning or anything of the kind. My intervention was a practical one; children didn't seem to understand what they were doing in these two areas and so they needed help. The multibased arithmetic blocks helped them to understand place value, because we varied the variables. I thought to myself, reasonably enough, that if you wanted to understand a mathematical concept you needed to vary the variables that were inherent in it; without such variation, you would only understand a particular case. To understand place value, you needed to have different bases. To understand algebraic manipulations you needed to have a large number of different embodiments of these manipulations in physical form. This is how the squares and strips and triangles and the balance beam, and paper cups and weights and all these sorts of things were introduced. Through all these different embodiments we tried to get children to understand what the distributive principle and what the commutative principle were; in other words, to comprehent at least partially what arithemtic and algebraic manipulations really were about.

Now when we got going with these activities, it very soon became apparent that we were playing with a very important mental process of learning abstractions, and so, from the practice of these school situations, I began to put together a very rough theoretical basis, in terms of which one could reconsider the whole of the practical situations and go on to further more refined experimentation. So in Building Up Mathematics, in Chapter 2 of the original edition, a theory of learning mathematics was proposed. It isn't perhaps worthy of the name "theory" but there was nothing much better at that time. Of course, this "theory" did not give us any detailed idea of what happened while children were learning.

This was the time when Jerome Bruner got to know about my activities. He read Building up Mathematics when it was first published, and thought it would be very interesting to examine the processes I had called into being, in a psychological laboratory. So he invited me to come to Harvard.

During 1960-61 Bruner and I worked on this problem. All we managed to do in that time was to generate a large number of sequences of learning situations, which we recorded, including everything the children said and did and felt, including their conversations and their overt emotional attitudes. We collected reams and reams of data. Bruner and myself were not in agreement about how to interpret the data, so we decided that I would write up my report and he would write up his. I wrote up my report, "An experimental study of mathematics learning" to which he wrote the preface. He later on wrote up his report in the form of a paper called "Toward a theory of instruction" which no doubt is well known. At the end of this, all we could really say was that there were some hypotheses about the kinds of process that might go on while the child was learning mathematics. The processes that needed investigation were abstraction, generalization, particularization, the relation between concrete and abstract, the way children symbolize, the way they interpret, in what order these come. All this was in a very vague state of affairs and it needed putting together in some more comprehensive form.

It was at this stage that I accepted the post of Associate Professor of Psychology at Adelaide University, where Professor Jeeves was the professor. We had worked together beforehand in Leicester on these problems on the theoretical plane. We had run between us an interdisciplinary seminar for a couple of years in which statisticians, philosophers, electronic scientists, psychologists, and mathematicians participated. We just brainstormed and talked around the problems. At the end of the Harvard year, Jeeves finally arranged for me to go over to Adelaide.

In Adelaide, two things needed to be done. First of all, I needed to persuade the school system that something after the fashion of the Leicester situation could be established so that we could have something to look at. Secondly we needed to look at some very detailed mathematical learning situations which we could put under the psychological microscope to see what happened from minute to minute.

This minute to minute description of what happens is described in "Thinking in Structures," and the book that follows: "The effects of structural relations on transfer." This was the result of our theoretical work. The practical situation was also developed at the same time. We brought in logic, sets, relations, at this point, using the Leicester type of "experimental" learning situations.

Of course, all this had to be written up. This was done in a book for MacMillan, Mathematics in the Primary School, which more or less described the situation as it was in the Adelaide days.

One thing that is important to remember is that Adelaide is at the "other end of the world". When you've reached South Australia, there's nowhere else to go but the South Pole. If you want to go anywhere from South Australia, on the whole it means a round-the-world trip. Once in America, it's even cheaper sometimes, to come back via Europe. Or if you go to Europe, you might as well come back through the United States or Canada. So it meant that every year while I was in Adelaide, I went on a round-the-world trip. Some of the things we were doing in Adelaide thus became a little better known.

Residents of Australia play the "hemisphere game" between Australia and the Northern Hemisphere because summer is replaced by winter and winter by summer as you change hemispheres. When there was nothing to do because it was too hot in Australia, I found myself in the Northern Hemisphere. This was one of the mechanisms under which the situation improved.

During these round-the-world trips, I met a number of people who began to show an interest. I had met Bob Davis already during the Harvard year, the initiator of the Madison Project. I had also met Max Beberman and a number of his co-workers such as Jack Easley who had worked in Hilo. He had been working in Illinois with Max Beberman and so was instrumental in inviting me to the Illinois Mathematics Project. I also met David Page, who had also worked at Illinois. I cemented these contacts, every time I went on an around the world trip. I gave progress reports of what we were doing in Adelaide, in Hawaii, in New Guinea, and various other places where I left some traces.

One of these places was Hungary. In 1962, the Second International Congress of Mathematicians was held in Budapest. I made a speech there about mathematics teaching which pushed certain people into action. To cut a long story short, Varga Tamás and his associates took up the ideas and after some difficulties, established some project-schools. These eventually grew into a nation-wide project. Every time I would to to Europe, I would also make a call in Budapest, and feed them some new ideas, some of which they incorporated. The joint work continued in this way.

Budapest was not the only place of ferment. There were others . We started a project in Ravenna with the cooperation of Angelo Pescarini, who was very interested. He worked up a situation in Ravenna and its environs where similar mathematical learning situations were created.

I also visited England a number of times. I made a number of appearances in France, in Paris particularly, and in Germany. In 1963 I went to Berlin and almost every year since then I have given one or more workshops in one or the other of the German cities. These workshops had been worked up to a kind of a standard format where one introduced some ideas and then, with the aid of local children from the village, everyone became involved in mathematics learning through the manipulation of a concrete situation. At first the teachers would be observing. After the children had left, the teachers would discuss what the children had done. This would be repeated every day, and the teachers would also join in the activities which they had observed the children do; after these activities of observation and manipulation, would come discussions about the underlying principles, psychological and pedagogical.

In this way, I was able to influence considerably what was to happen in Germany and perhaps in a number of other European countries. At one particular visit I made to Sherbrooke at the instigation of some of my French colleagues I gave a two-day seminar helped by second grade children acting as guinea pigs. This was so successful that they wanted me to come back for another two-week session the following year. I did this two-week course and the University of Sherbrooke as well as the other universities in the province became interested and discussions were held with a view to having me come to Quebec. Eventually I decided that Sherbrooke was a good place because there were already some experimental classes going which I had established myself, which the teachers involved didn't want to discontinue after my departure. These first classes were nurtured through the post and through further flying visits. To cut a long story short, I found myself eventually in Sherbrooke, where, after a few months of stay, we established the Center for Psycho-Mathematics. The brief of this Center is to study explicitly the learning of abstract structures, linguistic structures and in particular the structures of mathematics. We have tried to bring in, not only language but also the fine arts, and we now have plans for introducing work in music so that mathematics, language, art and music might be taught as a somehow connected set of disciplines.

This is where we are now. I might add to this that through all these contacts and joint work with different colleagues, we were able to pull an international team together into some kind of a group which we christened the International Study Group for Mathematics Learning. After a number of years of years this became incorporated as a non-profit organization here in Canada. We have established regular mutual cooperation and meetings with each other and we plan cooperative research projects, of a phychological nature, evaluation projects, and also pedagogical projects. From here in Sherbrooke we try to radiate to all the other centers that are working in conjunction with us, but, of course, in no way slavishly dependent on us. We try to pool information as far as possible and help each other in this way. This is roughly the history of the background of this Centre.

Klein: Well, thank you very much. Just one or two questions. How is the Sherbrooke Mathematics Programme built up ? Is it a definite mathematics programme which begins at grade 1 and is systematically built up until grade 8, or has it passed through a lot of changes throughout the years?

Dienes: Oh certainly, it undergoes changes every year. It is quite obvious that the programme changes almost every year. This is a research centre and not a programme construction centre; that is, our job is not to construct a mathematics programme but to find out facts about how mathematics and other related subjects are learned. If we find some interesting new ways of doing things, naturally, we like to introduce them into classes.

People that come year after year to visit us always see something different. There is a kind of core which has remained more or less the same. We do base the mathematics programme on the conventional mathematical build up. This does not mean to say that we shall always do so. By conventional build up, I mean that cardinal number is considered based on sets, the property of cardinality being a set-property, ordinality stemming from the study of order relations. Cardinals and ordinals are synthesized mathematically and psychologically into the natural number concept; the "one more" idea brings in the first arithmetical function, thus helping to cement the cardinal-ordinal mixture.

This is perfectly standard, from the point of view of mathematical interpretation. What is not standard is the way we do it. We try and utilize psychological principles which we have elicited from our practical work both in the laboratory and in the classroom. Some of these principles originated from the first days when we began work in Leicester with the following two principles:

1) Mathematical variability principle.
2) Perceptual variability principle.

Mathematical variability means that we vary the variables as far as possible to explain to the child what a mathematical concept is. This is part of the process of generalization; when you pass from particular to general, you have to be aware of all the variables and how much a variable can vary within a system that you are learning. Those were the germs of a theory; now we have some considerable experience on the generalization process.

In trying to put across the abstraction of algebra to children, we tried to employ the multiple embodiment principle or the perceptual variability principle, as the case may be. We imagined that abstractions were obtained as the communalities of various different kinds of experience, where the differences were ignored and the communalities were put together as the structure to be learned.

In order to do this, we have put the child through a number of different stages. The development of these stages is part of our theoretical build-up at the moment. We now distinguish six different stages:

1) initial free play,
2) the game stage, where the children are playing by the rules,
3) the comparison stage, where they compare one game to another.
4) the representation stage, where we would pass on to a more abstract formulation of all the games of the same structure that a child has learned,
5) the symbolization stage, where we begin to establish a language and describe the abstraction the child has learned,
6) the formalization stage, where children learn about axioms, theorems, proofs, and the like, and eventually reach the axiomatic stage.

This build-up is again the result of our own evolution from the first stages in Leicester to the stage at which we are now, where we have this Centre, where we have the chance of investigating whether certain hypotheses in fact hold up in practical experience or not.

Klein: People often ask if there is such a thing as a Dienes method which is the same all over the world in different countries or whether there are different methods which can be put together because they have common features.

Dienes: I think there are some common features and there are of course some differences. As we were just saying, there are differences even from year to year here in Sherbrooke so there are bound to be differences from place to place. Different people have different interpretations; they put different accents on different parts of the learning, on different parts of the subject matter, and surely they also adapt to local conditions. What we do in Sherbrooke and what happens in a country school in England or in Italy or what happens in the bush in New Guinea or in an Australian slum would be very different, because you would need to adapt the learning to local conditions.

The essence of the method, if we can call it a method, is to make sure that a child's own experiences are the main source of his learning. Taking a symbolic funnel through which you mix up the knowledge and pour it into the brain of a child and then you test him and see if he gives the "right response" is the essence of the traditional method. The types of project that are evolved through the International Study Group attempt not to do this. It's very difficult to base everything on a child's own experience, in other words on discussion and on personal manipulation because it's difficult to train enough teachers in a short enough time for them to be able to handle the amount of responsibility that this entails. Teachers do what they can. Take for example the situation in New Guinea. It is very difficult to train a local teacher in enough knowledge or relations, of sets, of algebra, and so on, so that he can create and manipulate situations of the kind relevant to the child in the bush or on an island or in a fishing village, and at the same time would lead that child towards the abstractions that are necessary for him to learn if he is eventually going to lead that country. We have to get out of there one of these days, in fact very soon, and the local people who are left will have to run the place. In order to run a place, you have to know how to think in abstract terms; this is why we use mathematics as a training, to get people to think in abstracto; "in abstracto" meaning not in an ivory tower but meaning that you have a wider field of applicability. The more abstract a notion is, the more applicable it is. It's the abstraction of mathematics that has allowed humanity to make such enormous progress in technology, because it is the abstraction of it that has allowed the enormous number of different applications. Our emphasis on abstraction, our emphasis on generalization, our emphasis on group discussion and on the manipulation of the environment, are the corner stones; I think most of the projects that are working together with us do have, to a larger and smaller extent, these ingredients.

Klein: Now it is clear to me that the Dienes Programme is a flexible one, varying in time and space but still having some common principles which can be applied in most places and at different times. I wonder if we can describe some of the common objectives of this project which are found in different places all over the world. What are the objectives that you wish to reach with these children? What are the objectives especially that are different from the objectives of teaching mathematics in traditional ways?

Dienes: The objectives of traditional mathematics teaching are mathematics and the objectives of modern mathematics teaching are probably not mathematics or should not be. We must not forget that most children forget most of the mathematics they learn. Therefore, if our objectives were knowledge of mathematics we would all be failing 99.9% of the time, so if we want to keep our jobs, at least we need to pretend that we have different objectives. I think I could enumerate a few.

For instance, let's take generalization. Mathematics is general; the relationship between particular and general is a very important attribute of modern life. When you go from one situation to another in life, it is important to realize the things that are in common. Certain kinds of behavior are acceptable in certain countries but not in others because of some convention, but kindness is always excepted under any convention.

The reasons for a convention are our collective adherence to certain human values regarded in a general way. The generalization of a convention to the reason for the convention and the reason for certain types of behavior is desirable in order to facilitate human communication and human relationships. If one can understand this, then one has made a step towards a better world. We can establish a training ground for looking at what is particular and what is more general; this would certainly be one of our objectives. Whether we reach them or not, is another matter.

A mathematician is an inveterate generalizer. The first thing a mathematician will say when another mathematician tells him something is: "Oh yes, but I think it could be more general if you did so and so." In ordinary conversation, one does not do this. One does the opposite: "Oh yes, why did you do this differently. I did it like this." I think if this could be reversed, it would be certainly a great thing gained.

There are other objectives. For instance, the idea of coding. We code every time we speak. We decode everytime we listen, particularly when we have to speak somebody else's language or worse still, somebody else's ideology. Let's say a capitalist is speaking to a communist. They're probably talking about the same things. They probably have the same human goals, but they speak different languages and so they get angry with each other. If they could each decode what the other was saying and re-encode it into his own form of codification, there would be more chance of understanding. So I should imagine that the practice of encoding, decoding and transcoding, just as an exercise, could have an enormously important eventual effect on attitudes. In traditional mathematics the notations are always the same. For example, the X axis is always horizontal and the Y axis is always vertical. You tell a student to put the Y axis horizontally and the X axis somewhere else, and he will be thrown. He will not be able to do the problem. I think it is very important to use different notations, different ways of expressing things and to be able to use transcoding devices, translation devices, from one kind of situation to an equivalent kind of situation; or even to an unequivalent situation, to be able to do a partial trans-coding known as a homomorphism as opposed to an iso-morphism. At least a part of it can be coded even if not all of it. Let us understand part of what you are saying even if we cannot understand all of it. The coding and decoding efficiency is a very important objective which we could achieve.

Another objective is self confidence. I think, it's very important that children should be confident that they can do things. It is not so necessary that they should know particular things, because what we can teach them to-day may have absolutely nothing to do with what they may need to know in twenty years. Possibly the mathematics which they will need has not been discovered yet.

How then can we teach them ? What we can teach them is how to cope with a situation which is uncertain, how to face the unknown, how to be pleased by tackling a problem and regarding it as a challenge In other words to look at problems as part of life, part of existence waiting to be solved; we are not to be annoyed or belittled if we cannot cope with something. The minute a child says:"I can't cope with this," almost excludes the possibility of his coping with it. He should acquire confidence in himself and say: "Oh yes, here's a new problem; I'll cope." That's already half the battle. His self-concept should be such that he can say: "Oh yes, I can manage," instead of "Oh I haven't done this yet", and then just sit there and wait for someone else to help him or do it for him.

Another important objective could be to learn how to reason in a seriously logical way; to learn how to discover fallacies in other reasonings; how to identify premises, how to identify conclusions. Very often an argument does not even get off to a start because the protagonists have different premises without knowing it.

Discussions about identifying initial sets of premises, identifying the kind of arguments used, would seem important. Questions of the kind: "Is this allowable?" "If this is allowable, why is it allowable?" "If this is an error, why is it an error?" The analysis of logical reasoning should form a very important part of our objectives.

Klein: I would like to mention some of the objectives we derived from the Hungarian project, but which I think are common to the Sherbrooke project and in other Sherbrooke type projects.

One of these objectives was to decrease the rigidity of the students thinking. I think it certainly should be one of the objectives. The project tried to reach this objective by providing multiple embodiments of different ideas.

Dienes: Yes. Certainly. Increased flexibility of decreased rigidity should be, I agree, one of the rather important objectives. This can be reached perhaps not so much by providing multiple embodiments, although this is of course one way, but by inculcating an attitude in the teachers of accepting different solutions from the ones they had thought up by themselves. If a child comes up with a different solution or even suggests a different problem, he should be allowed it. If he says "Look, I'm not interested in the problem. Here is a problem that is much more interesting. I'd like to solve that. Look how much more interesting it is", then instead of saying "Oh no ! You do what I tell you to. You're the child", the teacher should say "Oh yes, indeed! How interesting. Tell me about it." In other words, if changing things is at a premium rather than the opposite, than I think that the child will learn to be flexible; otherwise he'll become rigid because what he'll think to himself is "Oh yes, the teacher has told me what to do. I shall do it just like that." Then, he will develop an attitude that everything is just so, because he has been told so. Eventually he will stop thinking. Certainly I agree that flexibility should be one of the important objectives.

Klein: Another objective that may be connected with flexibility is creativity. The Dienes-method puts at least as much emphasis on developing creativity in children than in developing intelligence. Naturally the latter should be a target as well.

Dienes: Now, this is one of the targets, but I think it is extremely difficult to get moving on this target with rather sophisticated forms of mathematics teaching. It means that the teachers, in order to allow for creativity, would have to know a lot of mathematics and they don't; and so it means that before we can really go for creativity in real mathematical thinking, we would have to have teachers themselves creative enough and mathematically knowledgeable enough to be able to see what the possible avenues are that the children can tread. I rather fancy that for the next five or ten years, as we try and introduce mathematical thinking, the creativity objective is not going to be very much in the forefront.

Klein: Is the situation better in Sherbrooke, where the schools have direct contact with the Center?

Dienes: Well, yes. I think that, from my casual observation of children, they certainly do come up with different things. They come up with different games. They invent different situations. But even there, you do have to encourage them. It is so much in their social system that the adults impose and the children obey that you have to keep telling them: "Look, I may tell you something, but you don't have to take any notice, you can do something else." They still do. In fact, if you say this to them, they begin to feel free and not so imposed upon and they probably would rather prefer to do what you have suggested. This exonerates them from having to think of something else. To be creative is intellectually tiring, So maybe we all have an escape mechanism against it. Let's face it;

out of thirty children how many are you going to get who are really creative? It still remains to be seen whether you can train for creativity. We don't know if we can because we don't have enough teachers who are able to give a sufficient amount of freedom to children. And we do not have enough school boards that allow it, because they will say: "You must cover the ground, children must know so and so by the time they are so and so old." Teachers are going to be worried that if they let the children "create", they'll not conform to the norms demanded by society. So there are all sorts of factors militating against creativity. Society itself militates against creativity. The moment society is formed, it has certain rules and regulations which keep society together. You cannot just do what you like because you have to consider your fellow beings. And whatever social system or society you live in there is bound to be a constraint towards conformity, and creativity is going to be just for the few, I rather fancy. This may be a council of despair but I really feel that it is going to be a long time before we have a society of creative people.

Klein: I quite agree, but I still think that the role of creativity is increasing and society needs more and more creative people.

Dienes: Yes, that is so.

Klein: And I also hope and expect that using this method children's creativity can be developed. I still look forward to it.

Dienes: Instead of one out of 30 we might get three or four out of 30. That would be quite good.

Klein: Another objective of this teaching method could be to decrease the anxiety of the students, the anxiety which can be fanned by the student's family and by the school. It is quite difficult to decrease the anxiety which originates in the family but I think it would be something at least to decrease the anxiety that originates in the school.

Dienes: I think this is a good objective, but I think it has to be qualified. I think there has been some research, showing that a zero level of anxiety doesn't encourage work; for example, you might just sit in a coconut grove. A certain level of anxiety is necessary for any work that has to be done.

Klein: Yes, I'm quite sure, but my observation is that the problem right now in most of the schools is that most children are anxious because they are afraid of the teacher.

Dienes: I agree. I think that the traditional situation gives rise to too much anxiety, and certainly if you walk into one of the classes that we've organized, you'll find it very much more relaxed. Our type of classroom organization will decrease anxiety, because children are allowed to stop what they are doing, have a little gossip and then go on. Very many times I've made such observations when I've given demonstration lessons. The "lessons" are not really lessons, but learning situations where the children are influenced by the environment. Some of these children would work for a quarter of an hour and then quit; having an intellectual overdose of the problem, they would go away or would talk about the baby that was born next door or something like that, and after three or four minutes without anybody saying anything they would go back to the problem. "Oh yes now, what was it we were saying ?", and they would attack the problem again with a new vigour without any interference from the teacher. It means that the lack of anxiety does make it so that they're allowed to have those intellectual "coffee breaks" and they don't have to be keyed up all the time in case the teacher is looking and use all sorts of defence mechanisms like looking busy, having a pencil in their hands, pretending to be writing something. All these defence activities make it virtually impossible for the teacher to know whether any work is being done or not in a conventially organized classroom.

In a relaxed classroom, you can tell straight away who is in trouble, because he doesn't have to hide it, and so the teacher can do diagnostic work in very much more effective ways than in a more regimented type of classroom which is run on anxiety as its motive force.

Klein: I think what you are saying has something to do with so-called achievement motivation; namely the type of achievement motivation which is supposed to be developed in children, originating from an intrinsic motive, which is to develop a kind of mood in the children to achieve something, not because it is imposed on them but because they feel the necessity to achieve it.

Dienes: Yes, and there is an awful lot of perseverence that comes into it. I've noticed a lot of children who will go on for over an hour trying to solve a problem and then when it's time to go to play, you almost have to make them or decide "OK, kids, if you really want to continue, you needn't play". I've even had the other day a couple of children spending the entire weekend working out a problem and coming back on Monday jubilant saying "Look, here it is. We've done it." This is the kind of motivation that we should aim at; it is what you were calling intrinsic motivation; it militates in favour of achievement but genuine achievement instead of a spurious kind of achievement when you must fill so many pages.

Klein: Finally, educators often speak about their objective to develop the full personality of the children; but in fact only a very few things have been done in practice. Do you think that your method can contribute something to develop the full personality?

Dienes: Well, it's a little difficult to know what the full personality is. I think you can say that our method has a very socializing influence. When you put a lot of blocks on the floor for a first grade child on the first day, they all grab because they all want them. They say "This is mine" and start throwing them at each other and there is pandemonium. Before many weeks are up, they have learned that it pays to work together, that it pays to share the materials and to make something together, for example a big pattern which is beautiful; one can use all the material, and it is pleasant to work with another person. You enjoy the mutual contact. This has the definite influence on socializing. In traditional teaching, there is very little socializing, in fact, there is no socialising that takes place, except in the playground. The children musn't talk, they musn't work with each other, because it is called copying and there is no possibility of cooperative work. In this sense, the method helps to develop the personality.

Klein: I think it is another important objective, that is to reach a healthy social structure of the group which is in this case the class, and it also does quite a lot to develop the students' personality.

Chapter 9

A Visitor looks at Sherbrooke

Shortly after his appointment as Director of the Research Centre in Psychomathematics at the University of Sherbrooke in Quebec, Canada, in 1967, Dr. Dienes visited some nearby elementary schools. He found their mathematics teaching traditional in both content and approach. The first of several schools to respond to his suggestions for change was the Ecole Eymard. A new mathematics programme was introduced and allowed to grow up with the children from the start of Kindergarten (age 6) to the end of grade 6 (age 13). When Professor Lant visited the school in the spring of 1973 he found the Dienes Project established throughout the school and recorded his impressions.

Impressions on Entering

The classroom door opens and after a 'good morning' from the teacher and children which sounds unmistakably genuine one is invited to walk round, hold conversations, or handle any equipment that has caught one's fancy. Visitors are clearly part of the game. One has not disturbed the peace by one's appearance. After a few minutes all is back to normal and the newcomer is left free to discover what is going on.

There are no tidy rows and columns of desks, just tables, little chairs and chests of drawers scattered about the room. Here is a little group of youngsters engaging in conversation that sounds very exciting. Over there is a little fellow checking something on the blackboard; he will return to his place in good time, perhaps stopping for a short chat on the way. Two little girls are coming up; they wish to pass to the far corner of the room where the cards and blocks for their next assignment are hidden in the fourth drawer down. The air is busy but nowhere does the volume of noise need cutting down (see figure 9-1).

What has happened to the teacher? A moment ago he rallied the class for the welcome but now he seems to have let business run its way. Ah, there he is, crouching down over some plastic figures and pointing out their arrangement to a little group. A few minutes later he will have moved to give a few directions somewhere else.

To the visitor whose employment is not in education all this may come as a shock. He will reflect on his childhood, when all worked on the same page of the textbook from 9 to 10, did only what they were told and quietly sighed with relief at the striking of the clock. In the Dienesian mathematics room the situation has changed. The overall purpose is to promote mental growth through the understanding of mathematics, and this seems to be a less simple matter than is commonly supposed. An hour's daily drill in rows

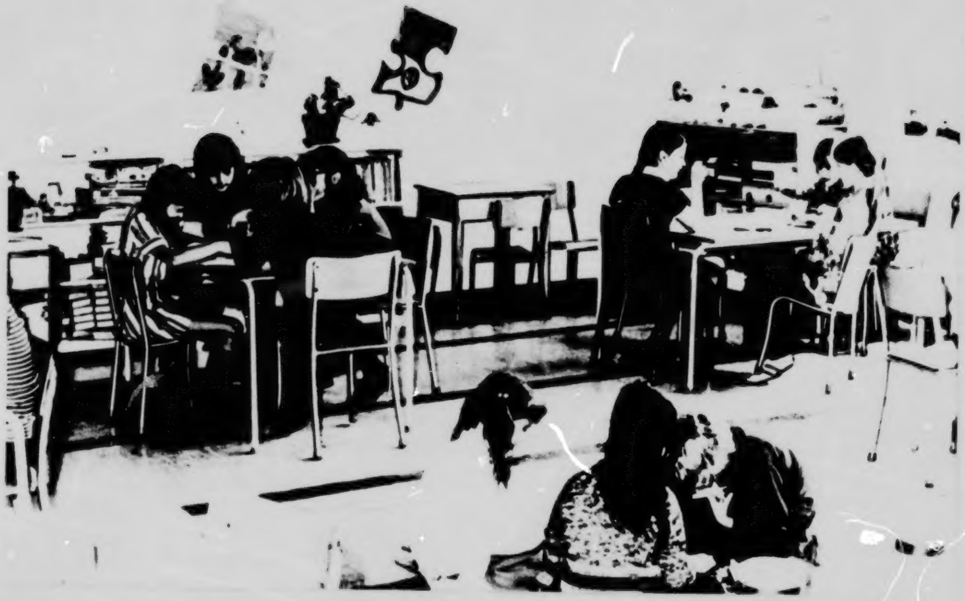


Fig. 9-1 There are no tidy rows or columns of desks.

and columns looks very tidy. The written work is easy to cross and tick, but the children are gradually dulled into a kind of automatic behaviour, totally divorced from the high-pitched thinking that has made mathematics the subject it is.

Should the visitor happen to be acquainted with today's trends in education then all this will seem far less heretical. Now and then the television informs the public of the silent revolution which hopes to build better and better schools for tomorrow. Psychologists and the social conscience are at last approving of what Pestalozzi, Froebel and Montessori preached some time ago: The environment of the classroom must be that which accelerates growth by providing for the child's needs at each stage of his development. Curiosity, movement and communication are necessary ingredients in successful learning, no matter what it is that is being learnt. And for this reason they are also fostered here.

A first look round.

Recovering from his initial surprise the visitor will most likely begin to look for the mathematics in which the children are engaged. Here a second shock lies in store. School mathematics is not what it was 20 or even 15 years ago. Back through our century learned societies often recommended changes in the kind of mathematics that was taught, but their voices rarely reached the elementary teacher before the Second World War. Following that event inventions in industry snowballed until it became clear that many employees would have to know more and more mathematics; the calculations composed of the basic rules of school arithmetic, it was argued, formed a vanishingly small part of what there was to know, and in any case proficiency in them was a less crucial matter with the coming of the computer.

On his departure the visitor will be convinced that arithmetic also forms a vanishingly small part of what young children can do. But for the moment he must trust his eyes when he sees charts, cards, blocks, tracks, arrows, and plastic men lying about on the work tables or the floor. The children are arranging, sketching, recording, disputing and agreeing, yet nowhere is there a single calculation with numbers to be seen. Mathematics without number, and yet mathematics in bewildering profusion! The only immediate indication that all might be well seems to come from the self-confident poise of those who are doing it all, and doing it as though they had never known anything different. And indeed it has all been set up to good purpose, as even a cursory glance at Dr. Dienes' writings will show. Viewed in the most general way mathematics is the study of relations between objects. What the objects and relations are does not matter so long as both can be described without any margin of uncertainty. The relation between a man and his dining club is clearly that of membership; we can always decide whether Mr. Jones does or does not belong to a particular club. On the other hand there is the relation between a man's life and the stars, as studied by astrology. Our knowledge of it is vague and of little use for starting any mathematics.

Our everyday environment contains many objects between which clearly defined relations exist: Looking at a row of trees, for example, we become aware of relations such as 'larger than', 'further than', or 'to the left of'. All of these play an important part in geometry and if we look in the right direction nature is prepared to present them to us in a simple way. More often on the other hand it is we who impose the relations on part of the environment we have selected: In chess or checkers every piece stands in a definite relationship to every other piece of the game. Rules in which these relations are clearly laid down form the beginning of a mathematical system, though in fact systems which usually bear this name differ from the games in the complexity of their rules, also in that the 'objects' which are manipulated are not nicely carved pieces but symbols like $\log_e 2$ or $f(x)dx$ written on paper.

Dr. Dienes has exploited this transition from the concrete to the abstract, from the simple to the complex, in an extensive programme designed to lead the children into the grasp of serious mathematics in a natural order. Most young children are far from abstractions and complexities, but already the first visit shows that their enthusiasm for detecting a multitude of relations among the cards, blocks, tracks and other equipment can be aroused and maintained.

Here lies a collection of cards. Each card has drawn on it a house with windows and chimneys. Some houses are blue, some red and some green. One house of each colour has 1 chimney and 3 windows, another 3 chimneys and 2 windows, yet a third 1 chimney and 1 window, and so on. Over there lie 4 kinds of geometric figures: Triangles, squares, rectangles, and circles. Each figure is moulded in 2 different thicknesses, 2 different sizes and in the 3 colours of red, blue, and yellow. Elsewhere again stands a table covered with a heap of plastic rectangles. These are embossed with tracks fitting together like the tracks of a model railway. At certain points holes appear into which stoppers can be inserted to act as signals for objects possessing certain properties to pass or not pass along the route.

Children who see the equipment for the first time view and feel it with excitment, like the contents of the toy box they find under the Christmas tree. Far from being regarded as a waste of time this initial appraisal is treated as a first stage involvement on which more purposeful activities can later be built. Just how much of this later development the visitor will be able to see depends on the grade of class to which he is first introduced, on the length of his stay, and clearly also on his rate of comprehending what is going on. The sooner he can supplement his direct impressions with some explanations the less fragmented and prone to misinterpretation will be the picture he takes home.

The children's personal relations

It may require several visits to the classroom before one can grasp the patterns in the children's activities. But whatever they are learning or not learning the children are persons, formed from one day to the next by the relationship they have with their classmates and the teacher (See Figures 9-2,9-3,9-4).

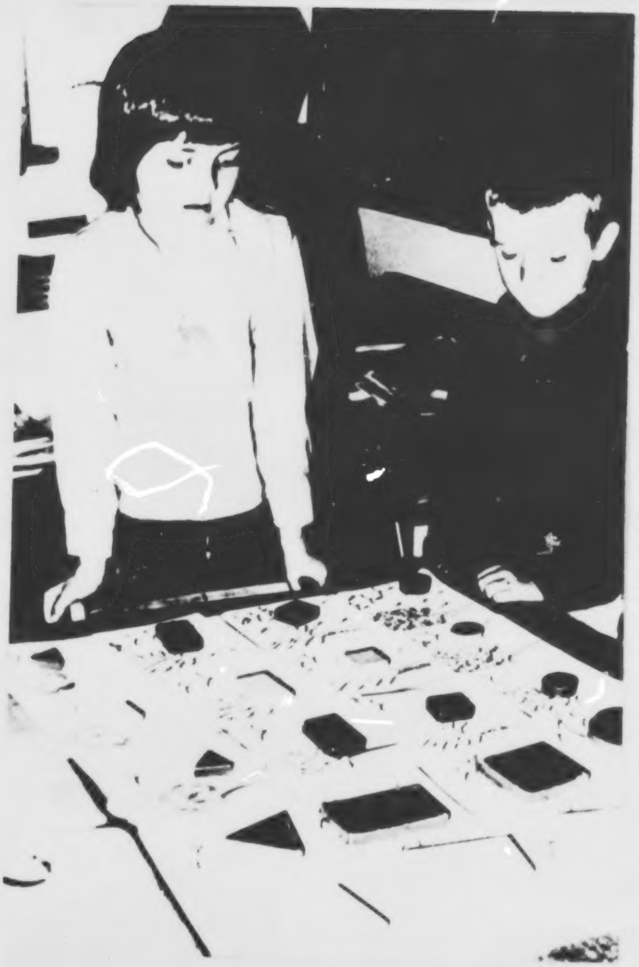


Fig. 9-2. Two ways of emodying the Cartesian product 2 x 3 x 3. Which block goes with which picture, and why ?

Fig. 9-3. Analysis after construction. Guess what I have changed !

Fig. 9-4 Mutuel discussion

Of this relationship the visitor soon becomes aware. Excepting the start and finish of an activity period or the occasional short demonstrations by the teacher no rigid discipline is imposed on movement, pace or speech. The children play and talk their way into mathematics, not away from it. If John moves over to his girl friend's table to look at what she is doing nothing will be said. He has taken a few minutes rest; before long he will be back in his group, busily arranging the cards with their houses and chimneys. Looking out of the window is Susan. Should she be recalled from a day-dream or will she come round herself? She seems very still, then suddenly walks to the blackboard with determination and alters a line or two in her sketch. She must have been weighing possibilities and suddenly made up her mind about what to do next. Maybe her friend Anne on the next table will reach the same decision more quickly when she takes her turn. It does not matter. The child's time is precious for what it does to him, not for what precisely he does in it. Accustomed to their liberty these children are far removed from "Oh, here's another lot of sums. Let's get down to the, through them and hope for the best!" To them mathematics is creation in which their whole being is involved. It may be creation of rules and creation within rules but because it is creation it brings with it contentment, perhaps a kind of aesthetic satisfaction which the research mathematician claims to experience in common with the painter, sculptor or musician.

The children are proud of what they are producing and ready to share their suggestions with each other and the teacher. If the visitor stands close by a table and looks as though he wanted to talk they will explain to him what they have discovered and what they are still trying to do. Suddenly a group in the middle of the room starts to buzz more loudly than the rest. It is not mutiny but high-pitched excitement over some unusual brain-teaser. No one punches his neighbour for being a fool or claims a monopoly in inspiration. Soon it all dies down because part or all of the answer is out, a joint production from which each worker has learnt his share. The better ones seem to teach the slower ones by example and the constant exchange of information. All have to practise clothing their thoughts in language until they are readily understood.

As the bell rings and the children line up to go out for their play the visitor will feel tempted to walk up to the teacher and after some congratulatory remarks ask him how he is able to keep track of his objectives with 30 children pulling here, there and everywhere. "I go to the groups one by one and when I think that a particular group has understood some part of the programme I sit down and give this group an evaluation. The rest carry on as usual." The answer is disarmingly simple. Until the crucial stage of formal evaluation is reached at the end of the scholastic year the terror of the mathematics test on Friday morning for which one crams the night before is gone; and when the year really draws to its close the children have enough knowledge and confidence behind them to put up a good show.

Other subjects are also there

The visitor has stayed for the children's return after 15 minutes. Walking round by himself he notices a table on which lie what appear to be children's drawings. Quite near to this stands another table covered with sheets of printed squares. There would be nothing unusual about this were it not that along the sides of these squares are printed words from the vernacular, ordinary words like 'sweet', 'light', 'bitter' and 'dry'.

Is this mathematics as well? If not, then what is the material doing on the tables during a mathematics lesson? Later he will ask for clarification but right now the workers are returning. Without fuss or regret about another spell of confinement they take up their places and continue. The group at the table with the squares cut them out and then match them up again in such a way that where a word and its antonym are given they come to rest against each other along adjacent sides. This is classification and grouping as in the case of the cards and plastic shapes, except that here it has a dual purpose: The 'objects' are words whose form and meaning make up the content of an English lesson as well.

It is all a little puzzling, and rightly so, if one considers the finer and finer specialization towards which occupations in civilized society are advancing. But then the society here is an embryonic one, where different priorities have been selected with profit. The claims of specialization are outweighed by considerations such as change in occupation, cost in the duplication of equipment, and the stress on co-operation and creation as underlying the gain of every form of knowledge. If the order to march in step has disappeared from the mathematics lesson why should it remain in force for the time-table as a whole? Art, language, geography and history can and do exist side by side with mathematics in the same room. Nobody seems to worry about unfair shares, about his subject catching some infection from shapes, arrows or hysterious codes.

Numbers too have survived

On his early walk round the classroom the visitor will as likely as not have been shocked by what the children did in the name of mathematics. Of course number somehow lurks in the background wherever people are busy with a multitude of objects, but not a single child seemed involved in giving number as such its due attention by working at sums.

Another teacher enters and invites him to look in next door. All seems much the same until he gets to a table where the children are handling some brown blocks and sticks of different sizes. The glamour of the lovely colours is missing but this is too superficial to be of serious interest. The visitor walks closer and notices that one boy has written down a row of figures and is just finishing the next one underneath. Arithmetic must still be alive!

How indeed could it be otherwise? New branches of mathematics it is true, are growing from year to year at an increasing rate; despite this an old and difficult branch, the theory of number, continues to be held in great esteem. Again, science and business call for longer and longer calculations to be done at greater and greater speeds, with the result that some mathematicians are leaving everything else to search for efficient ways of presenting such calculations to the electronic computer.

Far from ejecting number work from the classroom Dr. Dienes actually gave it close attention in his pioneering days some fifteen years ago. He produced wooden blocks to help children grasp the way in which we present our numbers by placing the numerals 0, 1, 2 .... in appropriate positions. In his equipment the number twenty-three, for example, is represented by three little cubes and two sticks known as LONGS. The latter have the same width and height as the little cubes but are ten times longer. Ten of the LONGS then make a FLAT which is both ten times longer and ten times wider than the small cube. Finally there is a fourth block to represent one thousand. This is ten times longer, ten times wider, and also ten times as high.

Organizing and understanding the representation of numbers is a habit which many children do not seem to acquire with ease. To sift the essential from the unessential Dr. Dienes went on to design collections of blocks based on organizations other than ten, namely on twos, threes, fours.. .. . .. right up to twelves. These provide great scope for activity and help children to surmount hurdles such as carrying. Paul, at whom the visitor is looking, has finished his addition sum as surely as he would have done in a classroom 30 years ago. But he seems to have worked with more self-assurance. On being questioned he reveals a good understanding of why he has carried, how much he has carried and where exactly he has written the number that had to be carried. The intricate convention about the position of the units, tens and hundreds has not been put before him in the ready-made state which man's thinking took millenia to evolve. Instead the blocks help him to retrace the invention in its growth; and the practice with different systems of blocks sharpens his awareness that position and consistency are more important than the choice of the number 10.

Dr. Dienes in the classroom

During one of his visits abroad a newspaper appearing just after another moon probe had left mother earth acclaimed Dr. Dienes as the 'father' of mathematics teaching. If the visitor is fortunate enough he will see him enter and take a look round in person. When this happens his standing as an innovator and theorist is forgotten; he just becomes the children's friend, a friend who has the knack of painlessly pushing their mathematical insight a little further ahead.

The heavy administrative commitments which match the extent of his travels and researches have never allowed Dr. Dienes to forget that the central reality in the whole project is the child, with his ayes and nays to what is tried on him. In the last resort teaching is a personal matter and the inspiration for striking out in a new direction derives from first hand experience more than from volumes of reports and statistics (See Fig. 9-5).

Dr. Dienes chats about, around and away from mathematics. The children enjoy the rapport, show off with the results of their mini-research and fall back on their regular occupation when he moves to another table. Eventually it is the teacher's turn for a tete-a-tete. He knows that his queries will get a patient hearing and be resolved. This done Dr. Dienes knows that now, if ever, he must sell the teacher another line, new sketches on a different number of cards, or just a sheet of instructions to try a subtle variation with material that is already known. If at this point the teacher reveals that he himself has tried something new so much the better. The spirit of meaningful co-operation extends to all alike. The worksheets handed to the children are liberally filled with "Try a different or similar combination yourself!", and the teacher needs no special invitation to try his hand at creation as well.

Fig. 9-5 Dr. Dienes in the classroom with five year olds.

To-day time and distance prevent many children working in the projects from seeing the 'father' in person. But those living near his home may even find themselves invited to supplementary schooling on a Sunday afternoon. They do not decline since they experience work and play as close partners at school. If the visitor has been asked for tea he may not find his host until it occurs to him to look in all directions on the floor. There he is at last, stretched out by the crouching child and waiting until each point of a finite geometry has been found its rightful place.

The preparation of the teacher

During and after tea there will be much for the visitor to ask and even more for which he will have no time to ask. How easy is it to find capable teachers willing to change over to the new approach which looks so demanding in its organization if in nothing else? This is part of a wider problem with which those responsible for the training of the elementary teacher are faced. On the one hand it is true that he can lead children into knowledge the more he knows about this or that subject himself; on the other there are strong reasons for letting him take the same class in a large number of subjects. As a rule he receives a broadly-based training which together with self-study or further courses during employment has to suffice. There is a growing tendency to lengthen his initial training and encourage him to advance to the standard of at least a general degree in 2 or 3 subjects. Despite these improvements a mathematics specialist in the elementary school remains a precious exception.

Every co-worker in the Dienesian project must somehow acquire a thorough grasp of its principles. There are no make-shift rules on which he can get by, no artifices for hiding his lack of knowledge. After making much additional effort himself he must accept it that the children are likely to bring him questions to which he has no answer to hand.

Although few regular courses exist at universities or teacher training institutions Dr. Dienes has, through regular workshops and consultations in the classroom, mustered sufficient teachers with the right personal and intellectual qualities to conduct classes in countries as far apart as Hungary, Canada and New Guinea. It is difficult to conceive how anyone not wholly in sympathy with the approach could induce children to receive its message. The thoughts of adults are more complex and more surely established than those of the young. For this reason they constitute a conscious or unconscious resistance to innovation, especially if this innovation is fraught with the hazards of humiliation. The children in their Kindergarten are starting with a clean slate, but the teacher has more likely than not been at the learner's end of traditional mathematics for some years. In one sense he must be untaught and start again; in another he must make himself director of a sizeable project and acquire such mastery over its details that he can admit any visitor with ease. When somewhere in the distant future the first teacher appears who has passed through the system himself we can expect its adoption to spread at a faster rate.

Taking a closer look

The first day's visit and discussions are over. There are many clear and good impressions but the visitor feels swamped by a mass of details which he can not properly connect. If he can spare a further 2 or 3 days he must start a more intensive tour in a Kindergarten or Grade I class, little by little making his way to higher levels from there. Even this may not help him over his confusion unless he settles on a definite item of instructional equipment and confines himself to tracing the uses to which this can be put.

Attention has here been focused on the equipment* known as trimath ('tri' as in tri-angle). This consists of plastic triangles in the six colours of red, green, yellow, black, white and blue. Each triangle appears with 0, 1 and 2 holes, also with 0, 1 and 2 vertices cut away. Generally it is immaterial whether the children' think of the shapes as triangles with vertices removed, or as polygons whose vertices number 3, 4 and 5. In the interests of simplicity they refer to them as triangles, hats and kites. There are 54 pieces in all (3 choices of vertex, 3 choices of hole and 6 choices of colour), but in many investigations only a small number of them is used. Considerations of space have made it advisable to change the colours to 3 in this account, namely white (w), grey (g) and black(b). This leaves 27 pieces in all (see Fig. 9-6).

---

* All the Dienes equipment is listed in the current catalogue of the Educational Supply Association, Harlow, Essex, England.

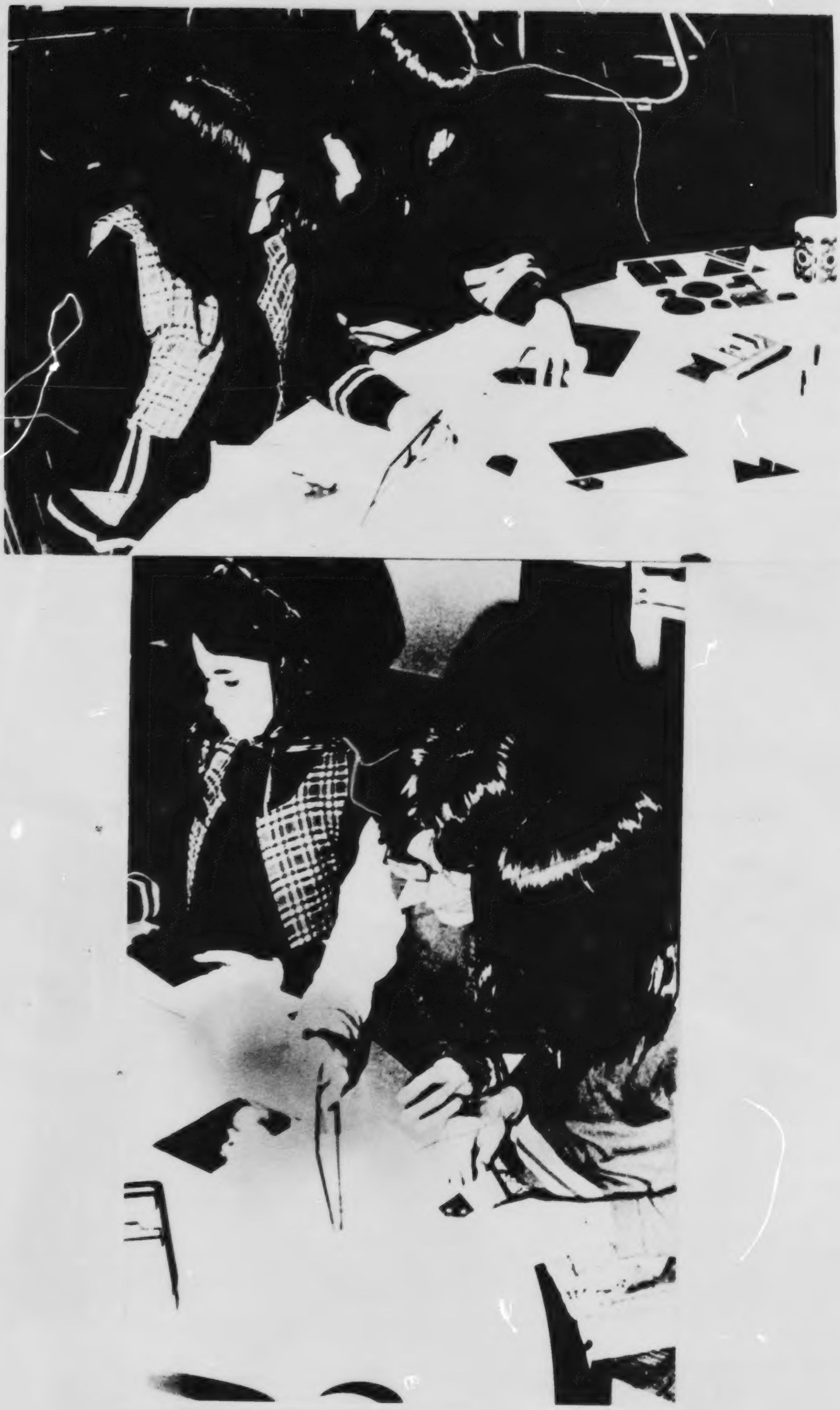


Fig. 9-6. Children working with Trimath and mirrors.

(i) A situation of free play

Already during the first day's visit it was noticed that children will tend to look at a piece of new equipment and begin to handle it rather like a Christmas toy. This kind of situation the visitor will most frequently encounter in the lower classes of the school. It is the one time that there is nothing for the teacher to prepare, but much to observe. Little Francis just delights in spreading the pieces out haphazardly, but next to him sits Susan, six months younger but more ambitious in her dealings with the triangles, hats and kites. Will they fit against each other until they completely cover all of the table top? It is worth trying, perhaps with the triangles alone, and sharpens her discrimination in any case.

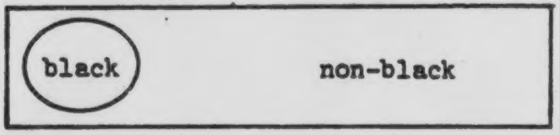
(ii) A game of familiarization

The children, it was noted, learn from working with each other. On this table they already have a pretty good idea of the kinds of pieces that the toy box contained. Taking the 9 white ones they appoint one of the group to do a little hiding while the others turn their heads away. When finished he challenges them to name the piece he has put away. The first one to shout "the kite with two holes" wins the round. There is a little scramble and then it starts all over again. Winners and losers have sharpened their sense for handling words.

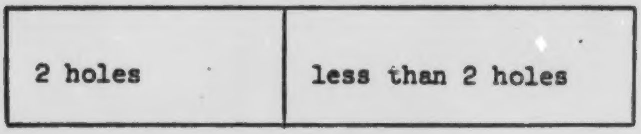
(iii) Starting some classification

In every science observation is followed by classification, the first attempt to impose our intelligence on the things with which we have to deal. Young children are no exception and the visitor pursuing the trimath will see pieces of white cardboard with charts like these:

(A)



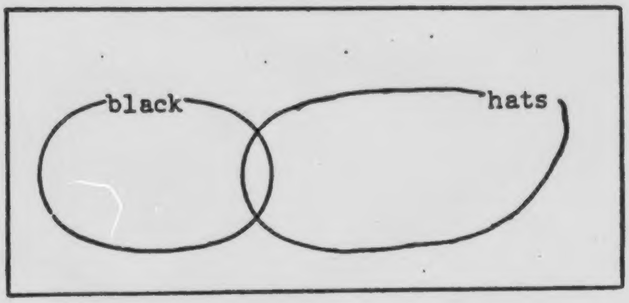
(B)



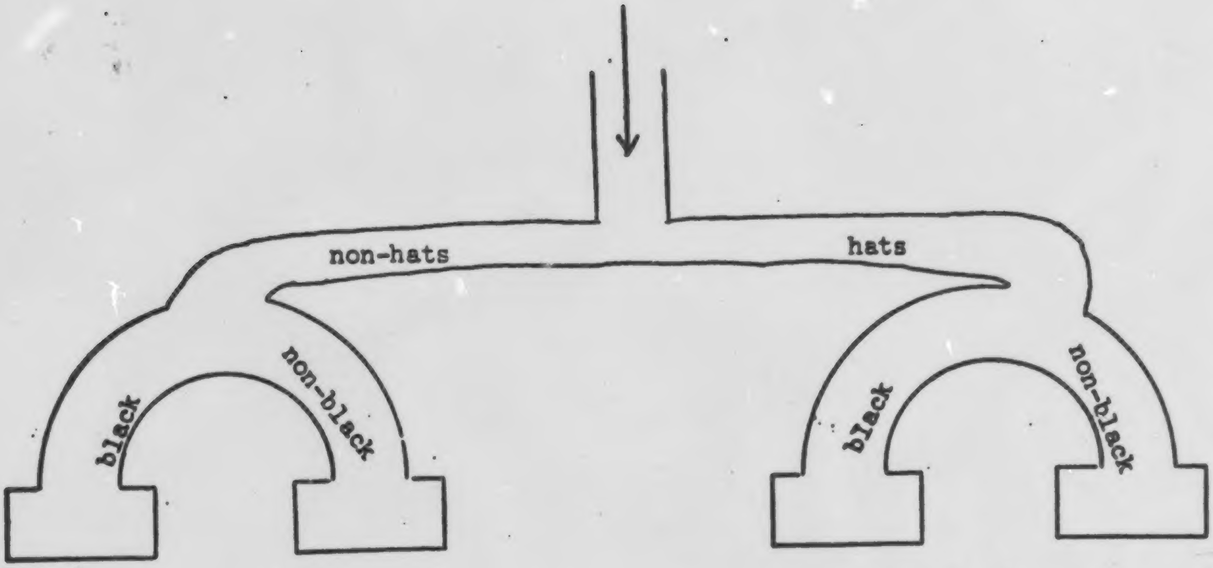
The children focus on just one property and then place the pieces in the appropriate region of the chart. If the full entire set is used there will be the colours grey and white to come under the heading of 'non-black', also both 0 and 1 as the number of holes less than 2. Should this prove too difficult then pieces with one of the properties are removed until more confidence is gained. For example, with the grey pieces removed then as far as task (A) is concerned the words 'white' and 'non-black' have identical meanings. The children may later check that this is so by interchanging the two words in instructions given to each other.

More difficult are the following:

(C)



(D)



(C) provides a first taste of the Venn diagram so commonly seen in school texts today, and (D) represents the same classification as (C) in what is often referred to as a decision tree diagram. To those children who have forged ahead and are waiting with everything neatly arranged the teacher will suddenly put some searching questions:

(1) Would the four different collections into which the pieces have been sorted remain the same if 'hats' and 'black' were interchanged in either diagram?

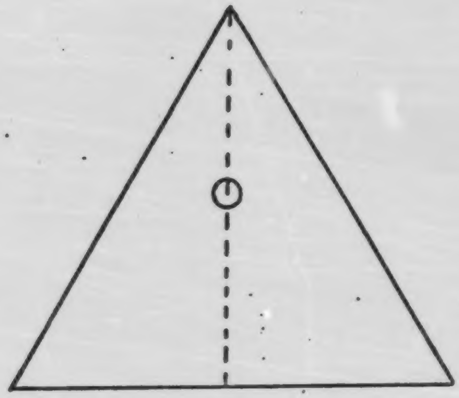
(2) Is there a change in diagram (C) corresponding to every change you can make in diagram (D), and vice-versa?

The second question is designed to test the child's power of abstraction, to what extent he has been able to free his mind from the particular shape of the diagram and grasped the essentials of the situation.

(iv) Trying some geometry

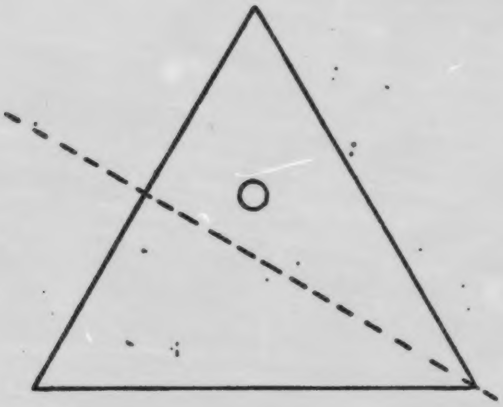
Even at the early stages of handling trimath some children will spontaneously have become aware that the shapes possess symmetry , some more than others. Their intuitive notions about this important property then undergo development through tasks such as these:

(A)



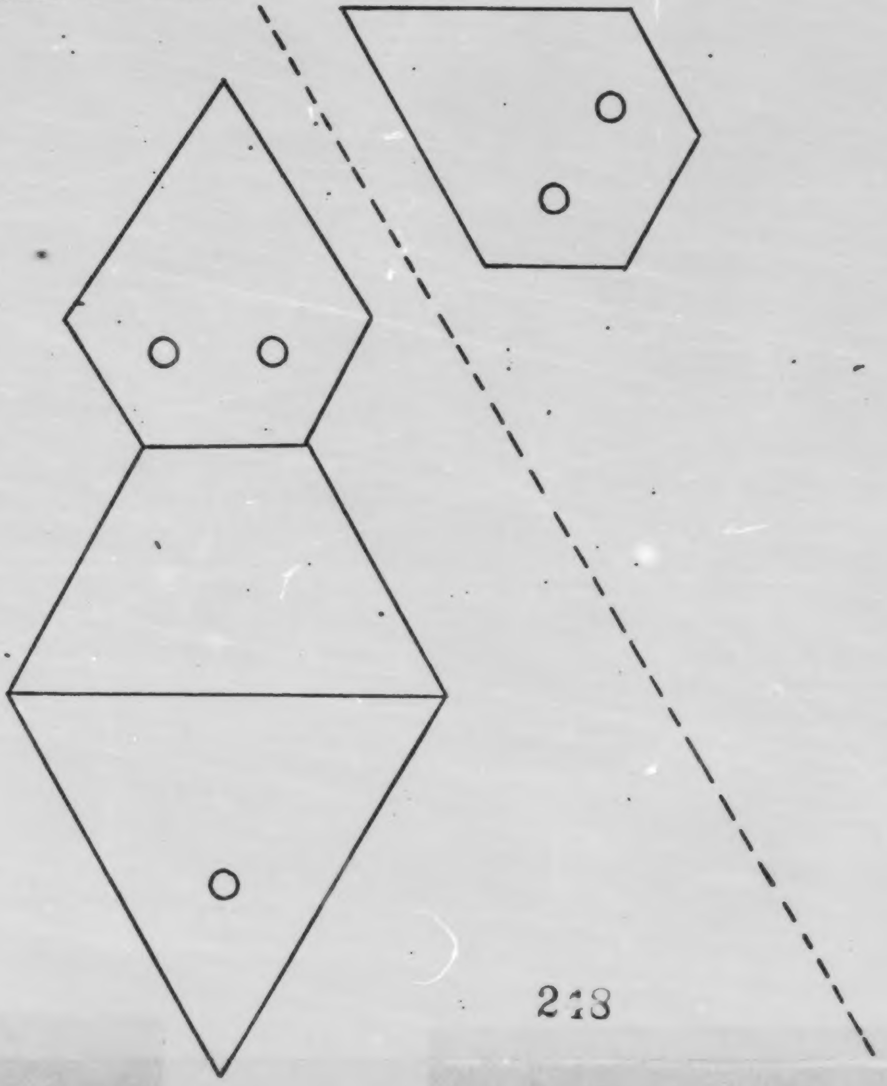
Take a mirror and place it along the dotted line of a triangle (any colour) with one hole. If you look in the mirror, you will see part of the figure reflected. Is this part the same as the part behind the mirror?

(B) Is the same true if you place the mirror along the dotted line shown here?

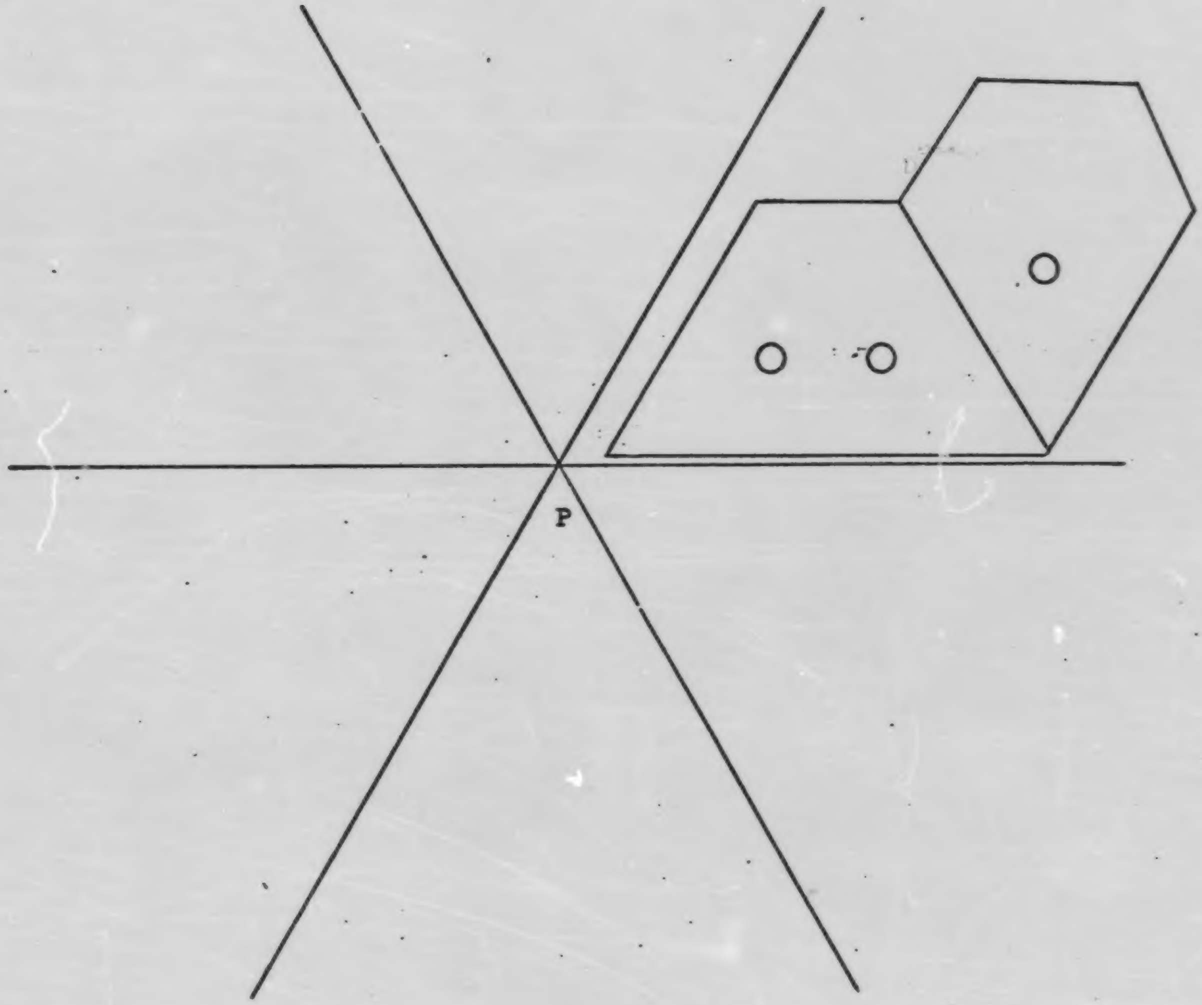


(C) Take some of the other shapes and find all their lines of symmetry.

(D) Complete the pattern on the right of the dotted line until it becomes symmetric about that line.



(E) Complete the following pattern until it has 60° rotational symmetry about its centre P.



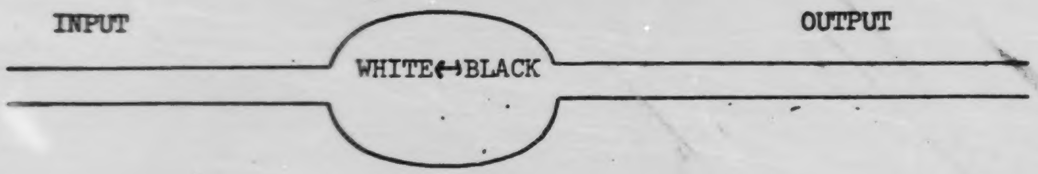
Change of colour destroys symmetry just as does a change in the number of holes. However, if a pattern requires many pieces of the same kind, the children will often agree to ignore colour. This trains them in accepting a somewhat unnatural convention and saves the inconvenience of breaking into several sets of the equipment on one table.

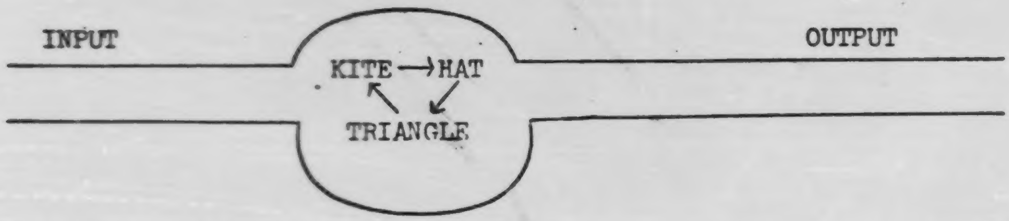
(v) Operators

An important feature distinguishing mathematics from other intellectual occupations is the use of operators. A machine which prints out enlargements of sketches represents such an operator, but the notion is quite general and applies in the abstract to multipliers, integrals or mappings.

For the children using trimath operators are commonly represented by drawings with verbal instructions on a large sheet of cardboard:

(A)





Operator (A) changes all white shapes into black ones and all black shapes into white ones if the output and input sides are interchanged. Grey pieces on the other hand it will reject, like a cigarette machine which has been built to take only certain types of coin.

Operator (B) will accept every type of piece. In one direction it will turn kites into hats, hats into triangles, and triangles into kites; if used the other way the kites become triangles, and so on. The order has changed from a clockwise to an anticlockwise one, but this was not noticeable in (A) where only two properties were exchanged.

Very elaborate networks can be constructed by chaining these and other operators together. This simulates the longer manipulations of traditional work and prepares the children to approach mathematical techniques intelligently.

Alternatively one finds these operators represented by plastic arrows of different colours. These can be moved about like the trimath pieces but contain no print. Their function is totally flexible and fixed by mutual agreement at the start of each experiment.

(vi) Order

Unlike our number system 1, 2,3. . . . ., in which 1 comes before 2 and 2 before 3 the design of the trimath pieces does not suggest any particular order in which they could be arranged. To explore this important concept the children are given or invent instructions such as:

(A) Triangles before hats, hats before kits,

(B) White before grey, grey before black,

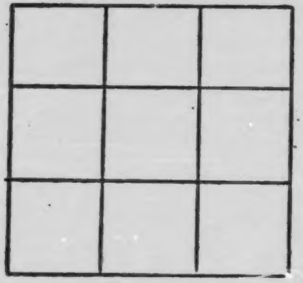
(C) 0 holes before 1 hole, 1 hole before 2 holes.

Even with these rules there is more than one way of setting the 27 pieces out in a row, since nothing has been said about the relative importance of (A), (B) and (C). To eliminate all ambiguity the children must add 2 further rules; for example, "(C) before (A) and (A) before (B)". To think of these is a more complicated affair since they are rules about rules, not rules about the pieces. If the children succeed they will have progressed to a second level of abstraction, a vital step in their preparation for detecting what is often concealed in traditional algebra.

(vii) Difference and similarity

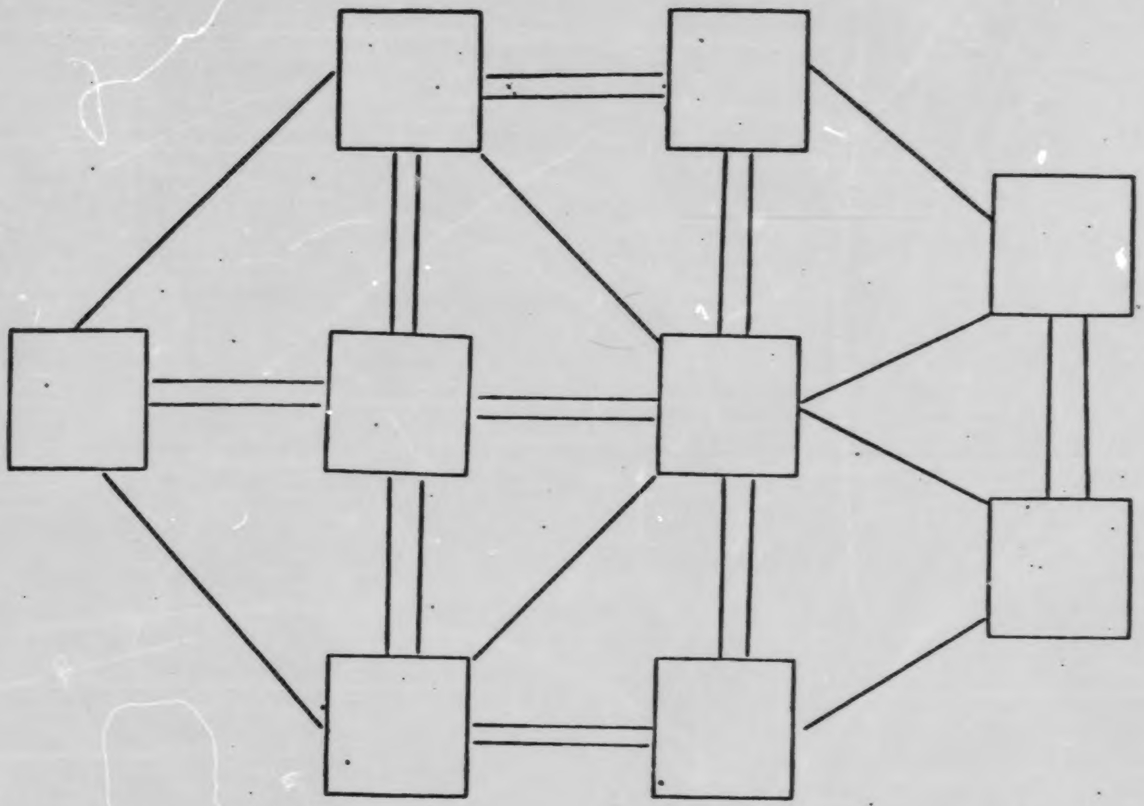
The process of classification is based on the fundamental relations of difference and similarity. Since any two trimath pieces differ from each in 1, 2 or 3 attributes a large range of exercises is open:

(A)



Use the triangles only and arrange them in the square so that two with the same colour or the same number of holes are next to each other.

(B).



Using the triangles, hats and kites without holes place them in the squares so that those joined by a single line differ in one attribute only and those joined by a double line differ in both colour and shape.

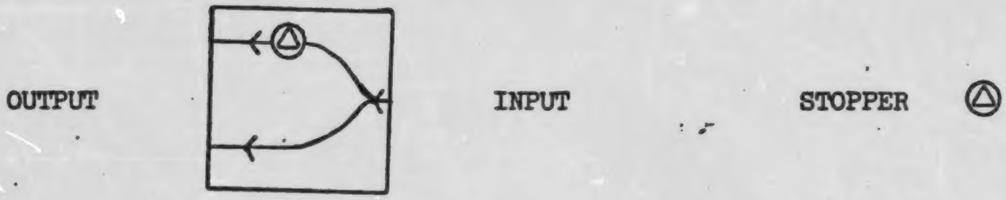
(C) The triangle has 3 vertices, the hat 4 and the kite 5. Each piece appears with 0, 1 or 2 holes. Add the number of holes to the number of vertices and you will obtain one of the numbers 3, 4, 5, 6, 7. Taking all the 27 pieces place those with the same number into a set (an equivalence class). How many different sets do you obtain and how many pieces are there in each set? (More sets can be obtained by assigning the numbers 0, 1, 2 to the three colours as well.)

(viii) The logic tracks

In mathematical situations one is often faced with an end result which has to be related to given information by several steps of strict logical reasoning. For children a concrete aid to such stepwise thinking is provided by the logic tracks. Each piece contains branchings of the track and there are also holes into which stoppers giving instructions on what is to go where are inserted.

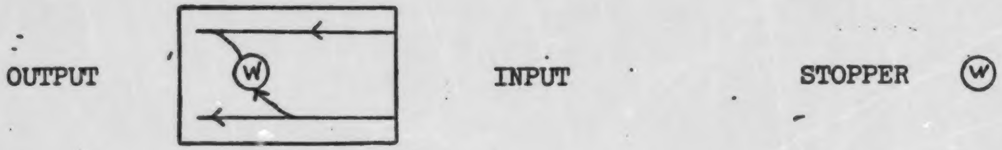
For example:

(A)



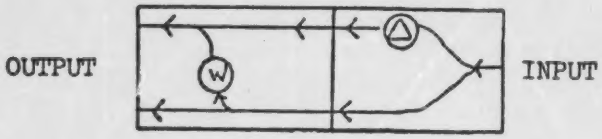
Encoded instruction: All and only triangular pieces must take the upper track, all other pieces the lower one.

(B)



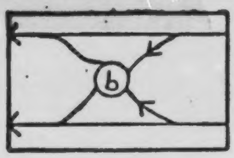
Encoded instruction: All the upper input goes to the upper output. From the lower input all white pieces are shunted to the upper output.

Suppose all the trimath pieces without holes are taken as input and tracks (A) and (B) are fitted together:



The upper output will now consist of all the pieces which are either triangular or white, or both triangular and white.

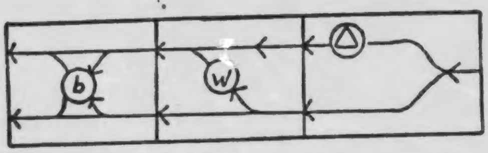
(c)



STOPPER ⓑ

Encoded instruction: All black pieces from the upper input are switched to the lower and all black pieces from the lower input to the upper.

If track (C) is fitted on to the combination of the other two we obtain:



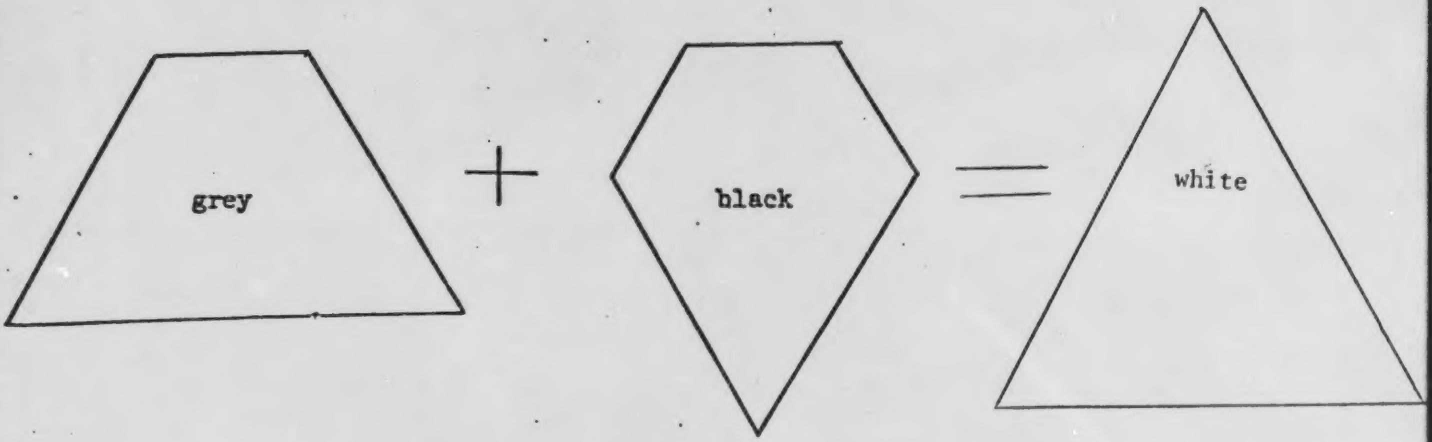
The upper output will now consist of all pieces which are grey triangles, or white; and it will also consist of all black pieces which are not triangles.

There are several other kinds of track. Taking various combinations of them makes it possible to put together a fairly complex set of instructions. Sometimes the children are given the final output and asked to find a combination of tracks leading to this output for themselves. If there are several possible combinations it is of serious interest to see which require the least number of tracks.

(ix) Little vector spaces

The reader will probably have felt that in the last three sections many of the descriptions taxed his thought processes more than did the earlier ones. This is because the visitor has been moving upwards and meeting with an increased concentration of mathematical content. Though still in the elementary school he has to work and ask his way around more and more. The hats, kites and triangles are still being arranged, but according to much more complex rules. Eventually

he stands puzzled before a blackboard on which is written:



This equation only has meaning if one first understands the following convention:

COLOUR		SHAPE		NUMBER OF HOLES
white	0	triangle	0	0
grey	1	hat	1	1
black	2	kite	2	2

Since each trimath piece has colour, shape and a number of holes it can be completely described by a triplet formed from some combination of 0, 1 and 2. For example:

The black kite with no holes is described by (2, 2, 0).

The grey hat with no holes is described by (1, 1, 0).

The convention further stipulates an addition of these triplets. In this only numbers in the same position are added to each other, just as real and imaginary parts are added separately in the addition of complex numbers. Also there are only three possible numbers for each position so that the answer to an addition must never exceed 2. This is easily arranged by working in a modulo or clock arithmetic. 3 is equivalent to 0, 4 equivalent to 1 and 5 equivalent to 2, just as a railway time table treats 00.00 hours as equivalent to 24.00 hours.

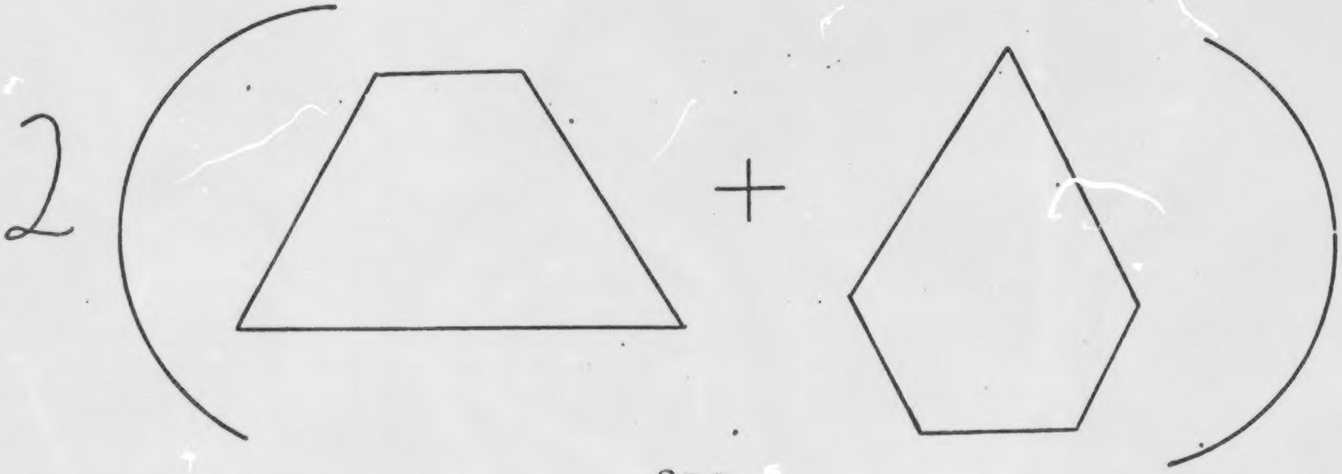
Written in triplets the equation now looks less mysterious:

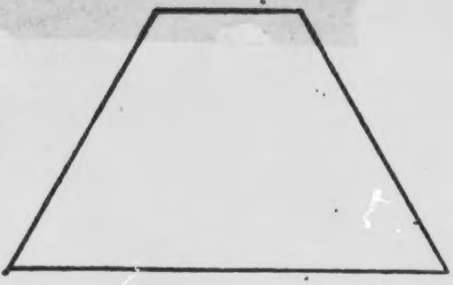
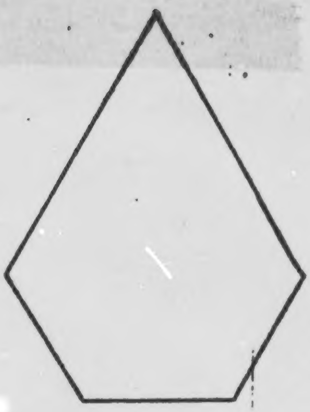
$$(2, 1, 0) + (1, 2, 0) = (0, 0, 0)$$

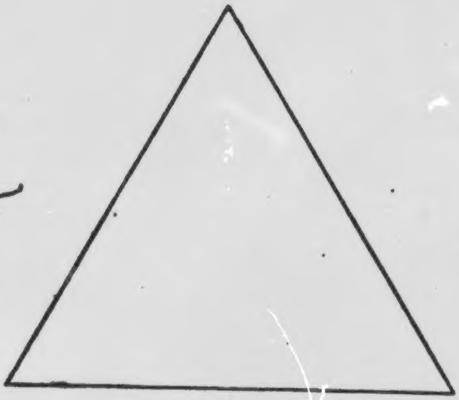
If it is also agreed that the triplets can be multiplied by any one of the numbers 0, 1 and 2 then the children are working in a system which is known to mathematicians as a vector space. An example of this multiplication would be:

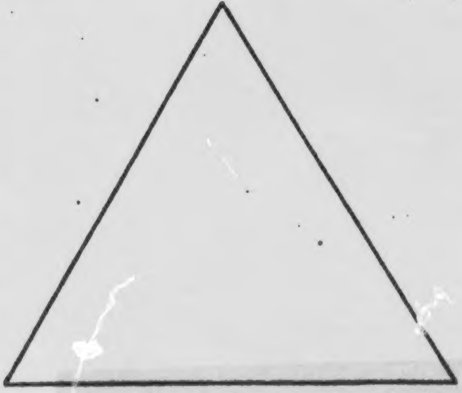
$2 \times (1, 2, 0) = (2 \times 1, 2 \times 2, 2 \times 0) = (2, 1, 0)$, since 4 is equivalent to 1.

With all this in mind the reader will be able to check that the following equation is both meaningful and correct:



= 2  + 2 

= 2 

= 

Much interesting work on subspaces, bases and transformations suggests itself. If from the 27 triplets the triplet (0, 0, 0) is excluded the remaining 26 may be interpreted as homogeneous co-ordinates for a projective geometry in the plane, with 13 points and 13 lines. From this again the removal of 1 line and hence 4 points will leave a 9 point affine geometry in which the missing line determines 4 directions.

Final Appraisal and Speculations

Here and now the new maths as developed by Dr. Dienes seems effective in the classroom. This simple fact speaks for itself, no matter how small the knowledge or how short the time the visitor has at his disposal. The children persevere without pressure, usually bring their tasks to a successful conclusion and often display unsuspected powers of originality in posing and answering problems of their own. There is no selection according to social status, no streaming according to ability. The approach is offered to the school population at large. Social and intellectual growth are fostered together in the democratic spirit which Dewey would have blessed. Of course one level of ability shows up against another in the speed, accuracy and creativity of the individuals who work and live together. But there is always something for everyone to learn, and to learn under conditions thought to be most conducive to his success. One hears of doctrines which simply divide men into those who can do mathematics and those who can not. If these need further corroboration they will not find it here.

To date materials and tasks have been developed as far as grade 6 and are undergoing trials in grades 7,8, and beyond. As far as one can judge the course is a comprehensive one. It brings children face to face with a very wide range of concepts as early as possible and ensures that their conceptual development advances towards the mathematics of the adult world on the broadest possible front. Against this gigantic network of organization one is tempted to view earlier efforts like those of Montessori and Stern as "special cases" in the slow revolution to present mathematics to children in a psychologically acceptable way. And the comprehensive nature of the undertaking also makes it difficult to conceive how the future could throw up a rival programme based on principles that were totally hostile or divorced.

What could be the long-term gains of working at mathematics in this way? Even if they retain no details the children will never accumulate that lasting sense of distaste for the subject to which even adults in intellectually demanding positions often confess. But the children's level of involvement is manifestly so intense that total obliteration seems highly unlikely. Instead we can expect them to retain a life-long curiosity and interest, no matter how small the mathematical demands their later occupations may make.

A more critical issue is the accumulation of knowledge and attitudes by those who will need mathematics as service subject in engineering, the sciences, or in the social field. If there is substance in recent criticisms that mathematics which is modern in either content or approach enfeebles proficiency in the traditional techniques the Dienesian programme takes care to accord these techniques their rightful place. More than that, they should be motivated by genuine understanding of the principles from which they derive. There is evidence to suggest that full understanding reduces the time needed for practice; add to this the manifestly more positive attitude to learning and the way into this or that piece of service mathematics should be straighter than before.

Finally there are those who will seek to advance to the frontiers of the subject in some area of mathematics, pure or applied. A small percentage among them is so highly endowed that they will advance to the top inspite of teachers, textbooks or courses. Immediately below them is a substantial number for whom the teaching strategies at school or college could make all the difference between failure and success. At the present time the Dienesian programme is just extending its feelers into the middle school. Conclusive evidence of its faults and merits may be lacking for many years to come. This need not stop us from raising two general questions to which mathematics education in the large might one day like to have an answer:

(1) Will the approach adopted here endow students at school with sufficient mathematical insight and expertise to render them immune to bad or indifferent pedagogy later on?

(2) Could the approach be extended with profit to all levels of mathematics learning?. In theory the principle of multiple embodiment as instanced in the wooden blocks or plastic shapes should facilitate understanding, no matter where it is applied. What must of course change is the type of embodiment. In higher learning concepts are generally abstracted not from blocks, colours or shapes within our everyday experience but from mathematical objects that are already highly abstracted themselves. Imagine a mathematician experienced in numerical analysis who wished to make some acquaintance with the highly abstract branch of modern mathematics known as category theory. What should

the introductory pages of a truly instructive text on this subject contain? To the categorist the ignorant reader stands in very much the same relation as the grade I child to his teacher. If there is a general solution lurking in the Dienesian approach then it might ultimately make communication among mathematicians less problematic than it seems to day.

Chapter 10

The New York Project

The OPI Project in Hungary and the Sherbrooke Project in Canada each operate within several schools, all of which belong to the public sector of education. There is no question of any social selection of the students, or the payment of fees. Generally students attend because the schools are conveniently situated near their home, not because their parents wish to send them there and are financially privileged to be able to make this choice.

Of the New York Project the exact opposite is true. It operates in a single, independent school which was founded for a section of society in which parents have both the desire and means to support an unusual venture. They sought to fuse the best in American and European educational tradition and out of this quest the Fleming School was born.

Historical Sketch of the Fleming School

The Fleming School, through which a Dienes-type programme has come to be established in the USA, was founded in 1957 with 30 students and 4 teachers in the confines of two rented rooms. It was to provide early education for children aged 3, 4 and 5, that is over the three years before the public sector could admit them to the first grade. The success of this early immersion was so startling that parents were reluctant to move their children when they reached the age of transfer at 6. Pressure to widen the school's range increased and a large programme of expansion was set in motion. Today the enrolment of the Fleming School is around 500 and all classes as far as grade 8 are housed in two large buildings, where space and equipment (Fig. 10-1) permit both teachers and students to enjoy the kind of educational setting that was planned. Those who leave after completing grade 8 have been known to obtain places in the best High Schools of Manhattan or in Preparatory Schools scattered over the country at large.

Fig. 10-1 The library of the Fleming school.

The Fleming School is chartered under the Board of Regents of the State of New York, but financially and otherwise independent of any School Board. It is a member of the New York Independent School Guild and maintains all necessary relations with external examining bodies and the public sector of education in general.

The Philosophy of the Fleming School

The outlook of the Fleming School is progressive, both as regards student-content and student-teacher relationships. It is held that the child is capable of benefiting from education in an institution at an early age, provided that the environment in which he is placed is structured in a way that will promote the growth of his personality as a whole.

The student-teacher relationships must be of the kind that will promote confidence and self-assurance, but at the same time also a spirit of mutual co-operation rather than selfish competition. Without being dogmatic or authoritarian in a conventional way teachers must seek to make the youngsters aware of the social constraints surrounding their existence. If this is done in the right manner the children will come to accept these restrictions as something natural and realise that the sacrifice of individualism brings rewards in its train.

The student-content relationships must be adapted to the stage of cognitive development which the young person has reached. Every child is curious, curious to learn more and more about his environment (Fig.10-2). In a progressive school this

Fig 10-2 Playing with the Dienes Blocks.

curiosity is fostered and enlisted as an essential constituent in the learning situation; it provides the dynamics of this situation by creating the motivation essential for the assimilation of content.

Stress is laid on discovery learning. The teacher provides a large collection of tasks whose complexity he regards suited to his students' age level. This provision takes the form of concrete materials and written instructions. As the child thinks about the instructions and materials he becomes aware of relations other than those to which his attention was directly drawn at the outset. This growth of awareness beyond that which is directly given is what constitutes discovery. The child finds new relations in the material if his mind is ready to find them. For this reason his discoveries are "natural" to him and will be more easily remembered than if they were just the discoveries ohhers, neatly packaged in some textbook language and presented to him in some unpalatable form of drill.

Stress is laid on creativity in both individualised work and group projects. There is creativity in art, music and dance programmes, creativity in the production of films and plays, and also creativity in the strictly intellectual pursuits of the science laboratory and the mathematics class. All creativity stems from a desire to produce. The desire to attain a chosen objective acts as a continuous motivator to search for the means of achievement, to examine what one has found, to discard the irrelevant and to use what is left as a means for moving nearer to one's goal. In the search for the means the teacher and fellow-students count among the resources that can be consulted (Fig. 10-3). This fuses the social with the intellectual aspects of the situation and goes a long way towards creating that overall harmony and happiness known as the "spirit" of the Fleming School.

Fig. 10-3 Discussing the peculiar aspects of a discovery

Bilingualism

The majority of non-American students at the Fleming School come from French families and the American students generally come from families whose experience of world travel has convinced them that knowledge of a foreign language is a valuable asset. These are among the considerations which have led to the introduction of a bilingual English-French programme commencing at age 3. Instruction is fairly evenly divided between the two languages and by the time a child reaches grade 3 his use of the French language comes quite natural to him.

Utilitarian considerations apart the study of a second language enriches the field of possible cognitive experience that is made available. If sentences written in two different languages express the same thoughts then they are, in Dr. Dienes' terminology, two embodiments of these thoughts, and the learner's changeover from one embodiment to the other facilitates his understanding of the thoughts for their true meaning. The practice of bilingual instruction, which is becoming more and more popular in the English-speaking schools of schools of the province of Quebec in Canada seems to be adding a new and important dimension to the curriculum, one that could not easily be provided by other means. Learning the second language as a side-show in parrot fashion is one thing, living out half one's school life with the second language as the chief means of communication quite another.

Teacher Selection and Preparation

In keeping with the cosmopolitan spirit underlying its foundation the Fleming School has also adopted a policy of recruiting its staff in those parts of the world where the best teachers of any specialism are to be found. The French portions of the curriculum are taught by teachers from France, and the subjects of English and Physical Education by teachers from England.

It seems generally agreed that the School's unconventional policy has produced satisfying results. Its approaches are rewarding but at the same time very demanding, especially on the teacher who has to break away from earlier habits and re-think his way through new aspects and new presentations of his subject from scratch. As a rule the children work in groups and each group works at its own rate. The teacher has to think of the needs of each group - what they are doing to-day, when they will finish, and what they will need to-morrow - and as far as possible also of the needs of each individual within the group. The faster learners need supplementary tasks of a more challenging nature, the slower ones repetition and many different presentations of what are essentially the same tasks. Only when a certain level of self-reliance and mutual trust has been established can he expect to go about his work unimpeded, and to establish this level he must in the first place have laid the foundation stones correctly.

One often hears it said to-day that with major changes in the attitudes towards authority the image of the teacher in the classroom also is no longer what it was before. In the co-operative atmosphere of a progressive school he can scarcely remain omnipotent and certainly not omniscient. The children no longer merely wait to be asked to repeat what he told them on the day or during the week before. Instead they come up with their own questions and expect more than an evasive answer. If the teacher does not know he must be prepared to admit this there and then, adding a promise that he will assist with the search for an answer in due course.

The Assessment of Progress

Sooner or later any programme which breaks with tradition faces the problem of meeting up with other streams of educational development at the point where it ends. During their years of attendance at the Fleming School students do not take public examinations Instead they are at regular intervals given tests of reading, achievement and intelligence. When these tests are corrected by the Educational Records Bureau their results are communicated to the School. Taken all in all these results, which reflect an extremely narrow sample of just cognitive variables, have never given any cause for concern. It must be stressed that they give no indication of the school's spirit and all its possible side effects, so that in any comprehensive evaluation they could only form a small component alongside the actual performance of leavers in their later education and beyond.

Initial Contact with the Dienes Method

In 1967 a teacher from the Fleming School read about the Dienes method and decided that its principles were in complete accord with the philosophy of his school. To make closer enquiries he decided to pay a visit to Dr. Dienes' headquarters at the Research Centre in Psychomathematics at the University of Sherbrooke. On his return to New York he convinced the school's director and several other teachers that a mathematics programme as developed in Sherbrooke would be ideally suited to their requirements. Even in mathematics it was possible to adopt a child-centred approach without jeopardising one's obligations to the intellectually demanding structures of the subject. In fact the new method advocated that the children face the challenges of logical thought from the earliest stage, and the means for getting them to do this could be tailored to the level of mental development they had reached.

Adoption of the Dienes' Method

In the following year (1968) the Fleming School and the Sherbrooke Centre signed a formal contract. In this contract provision was made for key teachers at the School to study existing programmes at some of the Sherbrooke schools and also to undergo some systematic training at the Centre themselves. Dr. Dienes in turn agreed to hold regular workshops for the teachers and to visit the Fleming School in person several times a year.

It did not take very long before the plans for co-operation between the two sides began to be translated into action. That the reaction of members of the staff at the school was one of unaffected frinedship and gratitude is shown in the following letter from the School's Director to Dr. Dienes:

"I must tell you how happy and enthusiastic our teachers were about the workshops at our school last Friday.

Miss W. and Mrs. L. are infinitely grateful to you and want to thank you for all your solicitude during their stay in Sherbrooke. They both came back radiant with enthusiasm about what they had seen and they are really filled with admiration for your work.

May I take this opportunity to remind you that you have so kindly promised to make a three-day visit to our School on November 18th,19th and 20th. We are all looking forward to this with great expectation and pleasure.

I believe Mr. F. had promised to get some record cards from your secretary to send to us by Mrs. L. or Miss W., but unfortunately forgot. I wonder if you could possibly forward some to us as I want to further the present enthusiasm of the teachers to the fullest extent."

It was an example of the frank rapport so characteristic of members co-operating in the various Dienes projects. Aware of Dr. Dienes' tight schedule of commitments here, there and everywhere the Director put in a reminder about the next workshop. Dr. Dienes replied re-assuringly and sent ahead sample questions on which his workshop would be based:

Rules for sequences.

Start with very simple sequences such as

& % % & % % & % % & % % & ....

and ask the children to go on.

Proceed to more complex rules such as

A B A B B A B A B B A B A B B A..........

or P Q R p p q q r r P P P Q Q Q R R R p p p p q.............

In some cases it is enough to know one element in the sequence in order to tell what the next element is going to be. In other cases it is necessary to know two or three elements.

With little children do not attempt to get them to verbalize the rules. If they can continue the sequence faultlessly without any explanation, it is enough.

With older children, from about 9 or 10 onwards, it is possible that they can formulate the rules precisely and apply them.

For example, the sequence

1, 2, 5, 6, 9, 10, 13...................

can be"verbalized" by saying

"If it is odd, add one, if it is even add three, to get the next element."

Children can play games with each other by giving each other more and more complex sequences, for which they have to guess each other's rules.

For example, a harder sequence would be

5, 8, 4, 2, 5, 8, 4, 2, 5, 8, 4, 2 ....

A rule for this sequence would be

"If the number is prime, add three; if it is not prime, divide by the smallest factor greater than 1".

We can also use sentences or words out of which to form sequences.

For example:

"The cat ate the mouse, the mouse was eaten by the cat, the cat ate the rat, the rat was eaten by the cat, the dog ate the mouse, the mouse was eaten by the dog,.............. "

What are the two missing sentences ?

What is the rule for writing the "next" sentence ?

In this short extract taken from what was actually sent there are challenge and surprise as one passes from one line to the next. Here and there some basic knowledge of numbers is presupposed but several of the questions can be answered from the thoughful analysis of knowledge that is much more general. An important objective of the method is reflected, namely the widening of the range of activities which teachers have come to associate with mathematics at the elementary level.

Some time later a parent only superficially acquainted with the theoretical backgroun to the new mathematics method but not slow to see what it was doing for his oldest child came out with this appraisal in a letter to Dr. Dienes:

"My own work has moved away from mathematics education; I'm with a firm doing different kinds of anti-poverty work - in welfare, health, employment and the like.

My oldest child, however, is now in the kindergarten at Fleming, playing with attribute blocks as if they were always part of her, I presume ! I haven't had much direct contact with Fleming, but through a series of circumstances I was asked to write an article about the math program to help Fleming parents better understand what is happening with their children. Apparently there is a lot of puzzlement on the part of parents, and there hasn't as yet been an opportunity to work out a real parents' education program.

The editors of the parents' paper are very interested in having something printed which will assist the parents. Hopefully, this article will do this and at the same time do justice to your program. Most of the ideas come from Building Up Mathematics, as you will plainly see.

I do hope it is satisfactory to you or that it can be made so with only minor changes. From the little I've heard I'm sure that more information for the parents is a good idea.

I look forward to seeing you when you come later on in November. The newspaper will come out after you are here, probably at the end of November, so it could be a good follow-up to your visit. I hope that Sherbrooke is suiting you and that your family is well."

It was the typical reaction to something that could be seen to be doing much good and yet remained strange and puzzling - like the first message sent by cable or the first carriage that moved under mechanical propulsion. In an independent school the parents are as a rule more intimately concerned with their children's progress. Dr. Dienes has never failed to realize that they are everywhere involved as a third party, even if they are not as regularly and readily available as the children and their teachers. He agreed to hold a parent's meeting at which questions of doubt and requests for elucidation could be put to him from the floor. Whereas it would be vain to expect that the transactions at these meetings reach all the parents who do not attend and leave all those who do attend convinced, such open discussions with the world beyond the classroom walls are an integral part of the service which the Sherbrooke Centre provides. It has been pointed out that the voices pronouncing on mathematics education today are many. There is nothing to prevent parents from being exposed to some of these more than to others. Not working in the Dienes Project themselves they then only have the occasional meeting and the general impressions of their children's reactions to assist them in forming a judgment.

Through this policy of give-and-take the Dienes and Fleming School programmes continued to identify their interests from year to year until in 1972 the new mathematics method was regarded as firmly enough established to be subjected to some formal evaluation.

To understand why educational changes are slow - much slower, than say, the conversion of a fleet of aeroplanes from turboprop to jet - one has to consider the amount of adaptation that is involved on the human side. The teachers have to discard old habits for new ones. Most of them are not trained in mathematics to a high level and must in this condition come to terms with the demands of a programme that is as intellectually demanding as it is profuse. Some will visit the Centre in Sherbrooke and then tell the rest what they saw and learned. All have to sit down and try their hand at a wide range of unfamiliar activities before they will feel confident to try them with the children. As the children are always present the teachers' familiarisation has for the most part to take place concurrently with the initial try-outs by the children. It may take months and even years before a teacher can claim to communicate the ideas efficiently, and this time will be lengthened if his teaching load consists of other subjects besides. The Fleming School differs from elementary schools in the public sector in that its teachers are specialists, that is, they take their subject from one class to the next instead of staying with the same class throughout the day. This has some drawbacks but also favoured the rate at which knowledge of the Dienes' programme could be assimilated. But on the side of deficit there is the flux in staff. Newcomers who are brought in to fill the gaps of those who have left must themselves go through the whole process of adaptation from the start.

On the children's side there are also difficulties. All but the youngest of them have started with mathematics in a certain way and already adopted some orientations of like or dislike towards the subject, also certain ways of thinking about its rules and remembering them. The new approach aims to provide them with greater stimulation and incentives to think about mathematics more efficiently, but this can not overnight give a child of 8 all the experiences he would have had if he had started with the new method at age 3 or 4. In the Fleming School as in most other schools the Dienes' approach tried to reach every grade as soon as possible after it was approved by the authorities concerned, but this does not alter the fact that the majority of those attending in 1968 never reaped its benefits to the full. Children, it is true, change quickly and have marvellous powers of compensation when it comes to making up for lost phases in development. But the wealth of potential experience in the Dienes programme is so large - even in the first few grades - that anything more than partial compensation must be regarded as an impossibility. Added to this we have the fact that in 1968 the teachers' level of competence was far from established and that there would be no one capable of distilling and condensing the essence of the programme's earlier stages for the children in the middle school.

It is significant that the evaluation projected late in 1972 and actually performed in 1973 concerned itself with students in grades 4 and 5, that is with students who were in grades K and 1 at the time of the Dienes Project's initiation. It was thought that these children had enjoyed exposure from a reasonably early stage and over a period of sufficient length to warrant the assumption that the effect of the new method on their mathematical outlook and competence would be substantial. Some 6 months before this evaluation Dr. Dienes summarised his impressions to the Director:

> " As a result of my last visit to the Fleming School let me make the following observations:
>
> 1. Although the standard of motivation and of general intellectual involvement is high in almost every case, the classroom organization still falls short of the desirable, mainly as far as the physical location of the working groups is concerned. This is due
> (i) to the fact that you have specialist teaching and so the classroom has to be re-arranged before and after each lesson, and
> (ii) to the fact that the size of the classrooms is simply not adapted to a very active situation, i.e. the rooms are much too small.
>
> Now that you are acquiring new space, would it not be possible to arrange things differently?

I suggest the establishment of as many maths classrooms (laboratories). as are necessary to house all those classes which have too little space. This would solve both the problems (i) and (ii). You would probably need a lab for every five or six classes.

Furniture could then be provided for storing the materials, instruction cards etc. in an easy filing system, tables could be used instead of desks and more blackboard and other feltboard space could be provided all round the room.

2. Although some of your teachers are now fairly knowledgeable in the maths involved, there is a great need for a full time maths co-ordinator. A person who has the mathematical knowledge as well as the pedagogical expertise to run such a program is extremely rare as yet. I had hoped to train such people at Sherbrooke. If we are lucky enough to obtain the services of such a person we must make his coming to work in New York worth while.

3. The evaluation will be done in February. This will include not only mathematical knowledge, but "learning ability" tests, a problem facing test, anxiety measures etc. so as to obtain as complete a profile of the school as possible, in comparison with a control school. Dr. K. is an extremely able psychologist who has already done such an evaluation for the Hungarian Project. He is working in Sherbrooke now and is about the only man in existence who has
   (i) the practical and theoretical psychological knowledge required
   (ii) an intimate knowledge of the practice, theory and underlying philosophy of a Sherbrooke type project to be able to give sensible answers to relevant questions. You will be getting all the information in January, about how to use and enrich our item bank for testing. I can discuss this with you on 20th December.

An Attempt at Evaluation

The evaluation just referred to was carried out during a 10 day period in April 1973, a little later than projected. It formed another member in a series of similar evaluations that had taken place in Hungary and in Sherbrooke.

Five questions were posed and answered on the results of tests given to students of all ability levels in grades 4 and 5.

1. In traditional mathematics students of the Fleming School were superior to those of a Dienes-type school in Sherbrooke.
2. In modern mathematics the same students were only marginally better.
3. Students at the Fleming School have a very positive attitude towards their teachers and their lessons.
4. The standard of learning ability is high compared with the school in Sherbrooke.
5. Taken all in all the attitude towards problems is less favorable at the Fleming School than might have been expected.

These results are no doubt in part attributable to the general atmosphere existing at the Fleming School. In no way do they seem to bear out the contentions of critics who maintain that a modern programme of the Dienes type enfeebles performance in mathematical skills of the traditional type. Had it been possible to administer these same tests to a traditionally orientated school of similar socio-economic standing in New York then the case might have been even more convincingly refuted. Arrangements for such a comparison were made but suddenly cancelled for reasons outside the experimenters' control.

The reader will appreciate that to obtain decisive answers in such matters is a lengthly, involved and costly business. On the informal side one could ask him to consider the satisfaction which both teachers and taught claim to have derived from the new method, also the high standards reached in external tests in algebra by those who leave after the completion of grade 8; they are often given a year's advance credit in this subject when they reach their high school class.

This evidence may do no more than convince the reader that the case is "not proven". If it does then let him rest assured that means for obtaining a more favourable verdict are still being sought.

The Situation To-day

Six years after its first exchange the contract between the Fleming School and the Sherbrooke Centre remains in force. Much of what was then a blueprint has become reality. Two specialist co-ordinators have been appointed and all classes now use or are strongly influenced by the Dienes approach. The three upper grades use a supplementary textbook in algebra from another author, but this does not mean that they have discarded the Dienes approach for what it offers in the way of thinking and living together. Although no compulsion is applied the efforts of two special co-ordinators for the method have made it possible for all teachers to participate and none have shown themselves reluctant to try the new materials and exercises as they reach the School. The transplant from Sherbrooke has here found a natural home, one in which even cursory inspection shows that all are happy to have been involved.

Some Misgivings

Parents continue to express misgivings about what is happening in the name of mathematics. Without going to the effort of a closer investigation they see no link between the games with coloured arrows, triangles and squares, and the calculations they have unconsciously come to identify with the whole of mathematics. Their misgivings have been endorsed by some poor performances in conventional computation tests in the middle section of the School, and by the hostility expressed towards the "New Math" in some recent literature (Kline, 1973).

The narrow identification of mathematics with numerical calculation is one which the Dienes approach does not and can not make. Number work has its important and rightful place at many points in the Dienes curriculum. But it is introduced as the expression of intelligent thinking rather than as conventional practice with large calculations. To meet the requirements of a traditional computation test at a stage set by fashion rather than any profound psychological considerations would entail a break with the carefully planned hierarchical structure that characterises the programme as a whole. Speed at routine calculations may be an objective of primary importance for a student about to take on employment in statistics or banking. This happens long after the date for which the tests here under consideration were fixed. There is some evidence that children who are trained over-early for this kind of performance tend to develop a narrow attitude towards numbers and sometimes come to regard them as somehow lying outside the scrutiny of logical thought. Unfortunately the Dienes method has not been long enough in operation to allow one to say categorically that students brought up throughout their school career in its stimulating atmosphere can learn such calculations easily and quickly when they know they are really required. Were this to become established as fact then of course all fears except those of the fashionable but ill-timed examination would be dispelled. At the moment the overapprehensive parent, it seems, must be asked to look at the method as a whole, at the positive contribution it makes to the attitudes, discussion and performance in the classroom. Many of the questions spontaneously posed by children would be totally meaningless to comparable children brought up in mathematics along traditional lines.

As regards the hostile literature this has generalised its venom against new mathematics as a whole because some new mathematics has been found unsatisfactory. Over the past 15 years many programmes constructed on widely diverse principles and overall objectives have sprung up throughout the world. All have been called new, but closer examination often shows that the attachment of this adjective is nothing more than equivocation as far as their inner structure is concerned. The Dienes programme is new, but there seems little justification for judging it on samples of students that have been taught some new method elsewhere. In any case, these attacks for the most part offer criticism without construction. To find dissatisfaction with the approaches of 30 or even 20 years ago one has only to look at the articles liberally scattered over the journals of mathematical associations in North America and in Europe. In the absence of any substantial body of positive suggestions the aims of today's neo-traditionalists must remain a matter for conjecture. Even if they reject the methodology as put forward in this book they have still to contend with the new branches of mathematics that are influencing the lives and activities of people who are not mathematical specialists. As an obvious example the human calculator is being replaced by the mechanical or electronic computer on an ever-increasing scale. This would seem to create a demand not so much for the speedy and accurate execution of routine calculations in one's head as for the ability to understand what these calculations are about, how the outputs of these machines may be estimated and checked.

BIBLIOGRAPHY

Ashlock, R.B. *Current Research in Elementary School Mathematics*, by Robert B. Ashlock and Wayne L. Herman publ. Macmillan Co., New York, 1970.

Biggs, J.B. *Anxiety, Motivation, and Primary School Mathematics: Occasional Publication No. 7.* London: Information Service of the National Foundation for Educational Research in England and Wales, 1962.

Cambridge Conference on School Mathematics, *Goals for School Mathematics.* Boston, Mass.: Educational Services, Houghton Mifflin Co. 1963.

Crittenden, B.S.(ed.) *Means and Ends in Education*. Toronto: Ontario Institute for Studies in Education, 1969.

Davis, R.B. The Madison Project's approach to a theory of instruction. *Journal of Research in Science Teaching, 2, 1964.*

Dienes, Z.P. The growth of mathematical concepts in children through experience. *Educational Research, 2*, 1959.

Dienes, Z.P. *Building Up Mathematics*. London: Hutchinson-Educational, 1960.

Dienes, Z.P. On abstraction and generalization. *Harvard Educational Review, 31, (3)*, 1961.

Dienes, Z.P. *The Power of Mathematics*. London: Hutchinson and Co., 1964

Dienes, Z.P. Mathematics through the Senses. London: NFER Publishing Co., 1973a.

Dienes, Z.P. Les six Etapes du Processus d'Apprentissage en Mathématique. Paris: O.C.D.L., 1973b.

Dienes, Z.P. On mathematical curriculum construction. Conceptos de Matematica, 1973c.

Dienes, Z.P. and Golding, E.W. Approach to Modern Mathematics. New York: McGraw-Hill, 1967.

Dienes, Z.P. and the International Study Group for Mathematics Learning. Mathematics in Primary Education. Hamburg: UNESCO Institute for Education, 1966.

Kline, Morris Why Johnny can't add. New York: St. Martin's Press 1973.

Lamon, Wm. Learning and the Nature of Mathematics. ed. William E. Lemon, publ. The Teachers College Press, New York 1972.

Lockard, J.D.(ed.) Eighth Report of the International Clearinghouse on Science and Mathematics Curricular Developments, The Science Teaching Center, University of Maryland, 1972.

Rosenbloom, P.C. Modern Viewpoints in the Curriculum. McGraw-Hill, New York, 1964.

Roughead, W.G. and Scandura, J.M. What is Learned in Mathematical Discovery. Journal of Educational Psychology, 59, 283-289, 1968.

Sherman, H.	*Common Elements in New Mathematics Programs*. publ. The Teachers College Press, New York, 1972.
Surányi, J., Reményi, K., and Varga, T.	*Principles of Updating School Mathematics as seen in Hungary*. Background paper of Working Conference on Mathematics Learning, Budapest, 1972. Manuscript.
Taylor, C.W.	All students are now educationally deprived. *Proceedings of the XVII. International Congress of Applied Psychology*, Brussels, 1971.
Varga, T.	A geometry lesson in a Hungarian school. *Bulletin of the International Study Group for Mathematics Learning*, 1, (3) 10-16, 1963.
Varga, T.	Pilot work in Budapest. *Bulletin of the International Study Group for Mathematics Learning*, 2 (4), 1-6, 1964.
Varga, T.	The use of a composite method for the mathematical education of young children. *Bulletin of the International Study Group for Mathematics Learning*, 3 (2) 1-9, 1965.
Varga, T.	Higher mathematics in lower grades. In: Feldman, L. (ed.) *Prepubescent Mathematical Learning*. San Jose, California: San Jose State College Mathematics Department, C18-C22, 1968.

VOLUME II

LEARNING MODERN MATHEMATICS - EVALUATION

by

Sándor Klein

Assistant Professor of Psychology

Teacher Training and Pedagogical Institute

Technical University of Budapest, Hungary

with the co-operation of

Ingrid Krause

Rsearch Assistant

Research Assistant in Psychology

University of Sherbrooke

Sherbrooke, Quebec, Canada

and

Rainer Lant

Associate Professor of Mathematics

Memorial University

St. John's, Newfoundland, Canada

Acknowledgements

To carry out a study such as this, one is always indebted to several persons. This is especially true in my case.

Concerning the Budapest study, I should like to express my thanks to E. Neményi for her work in the examination of hypotheses 1 and 2 (measurement of knowledge level), to Cs. Pléh for his great help with the examination of hypothesis 6 (creativity), to A. Kónya for her significant contributions to the treatment of hypotheses 7 (perseveration), 8 (children knowing each other better), 9 (attitude) and 10 (anxiety), and to L. Séra for his work in the examination of hypothesis 11 (achievement motivation). The instrument used in hypothesis 5 (measurements of learning ability) was prepared by Zs. Kámán. A large part of the work was performed by experimenters, Cs. L. Nagy, and my students, M. Kirják, I. Kovács, F. Lidi, L. Sárkány, and I. Standeisky, who chose this study as the subject of their final theses. Financial support was given by the Országos Pedagógiai Intézet (OPI: Hungarian Pedagogical Institute) and the Pedagogical Department of the Technical University of Budapest. The section of this book concerned with the Budapest study is largely based on a report (Klein, 1973), prepared with the help of S. Hennessey.

Concerning the Sherbrooke study, I would like to express my sincere thanks to Ingrid Krause, my main aid in preparing this part of the book; Sean Hennessey, Elizabeth Kratze, France Girard and Jean Gagné, who administered and scored the tests; Jean-Marie St. Pierre and Pierre Dion for the preparation of the computer programmes; Bernard Connes and

Pierre Colin for assistance in the statistical design; Graham Cantieni for the some of the photos; Madeleine Fortier, Diane Belley, Michèle Vallée and Marie-France Kyle for secretarial duties and some graphic work; Janice Belkaoui for correcting and editing the manuscript and John D. Williams for his continuing advice and encouragement. This investigation was carried out with the help of Canada Council Research Grant S70-0982. Financial support was also provided through the University of Sherbrooke, Quebec, Canada.

The New York evaluation was carried out primarily with the help of Ingrid Krause. Most of the people mentioned in connection with the Sherbrooke study also aided us with the New York material. Financial support for this study was provided by the Fleming School.

For all three studies, I wish to express my special thanks to the students, teachers, and directors of the schools involved in the work.

The entire book was prepared while the author was a guest researcher at the Centre for Research in Psycho-Mathematics, at the University of Sherbrooke, Quebec, Canada; completing work in Part II (Psychological Test Development and Evaluation) of the project: The Development of Tests for use in the Sherbrooke Programme. (For Part I of this project, see Williams, 1973, and Connes, 1973.)

Final editing and correcting work was made by Rainer Lant.

I would also like to express my appreciation to Z.P. Dienes for his continuous aid and encouragement in the one and one-half year long period spent on work in this field.

In spite of the great contribution of all the above-mentioned people, the author must take responsibility for mistakes in the planning, implementation and reporting of this research.

Sándor Klein.

Budapest, 1980.

CONTENTS

Page

Chapter 11: The nature of evaluation 1

11.1 The origins and latest developments in educational evaluation

11.2 Evaluation models

11.2.1 Definition and classification

11.2.2 Judgements versus decisions

11.2.3 Summative and formative evaluations

11.2.4 Informal evaluation

11.2.5 Formal evaluation

11.3 Common elements in modern evaluations

11.4 The necessity for improved evaluation methods

11.4.1 The need for innovation in teaching methods and objectives

11.4.2 The need for feedback through evaluation

Chapter 12: Difficulties in educational evaluations 15

12.1 Finding behavioural objectives

12.1.1 Aims

12.1.2 Objectives

12.2 Criteria and methods which should be considered in an evaluation programme

12.2.1 Super-abundance in evaluating dimensions and methods

12.2.2 Shortage of relevant measurements

Chapter 13: Theoretical model of the evaluation for the Dienes programmes 37

13.1 Objectives

13.2 Personality concept

13.3 Hypotheses

13.4 Statistical design

13.4.1 Population of interest

13.4.2 Selection of groups

Page

13.4.3 Variables of interest
13.4.4 The problem of measurement
13.4.5 Sampling
13.4.6 Statistics
13.4.7 Generalization

Chapter 14: The structure of the three evaluation studies 58
14.1 The Budapest study
14.1.1 Purpose of the study
14.1.2 Organization
14.2 The Sherbrooke study
14.2.1 Purpose of the study
14.2.2 Organization
14.3 The New York study
14.3.1 Purpose of the study
14.3.2 Organization

Chapter 15: Traditional and modern mathematics 73
15.1 Knowledge of traditional mathematics
15.1.1 Theoretical background
15.1.2 The Budapest study
15.1.3 The Sherbrooke study
15.1.4 The New York study
15.2 Knowledge of modern mathematics
15.2.1 The Budapest study
15.2.2 The Sherbrooke study
15.2.3 The New York study

Page

Chapter 16: Fatigue (Budapest study) 106

16.1 Treatment of data

16.2 Theoretical background and methods

16.2.1 Reaction time for visual stimulus

16.2.2 Reaction time for acoustic stimulus

16.2.3 Heart rate

16.2.4 Visual field

16.2.5 Finger tremor

16.3 Results

16.3.1 Reaction time for visual stimulus

16.3.2 Reaction time for acoustic stimulus

16.3.3 Heart rate

16.3.4 Visual field

16.3.5 Finger tremor

Chapter 17: Attention (Budapest study) 118

17.1 Theoretical background

17.2 Method

17.3 Treatment of data

17.4 Expectation

17.5 Results

17.5.1 Discussion of the results

Chapter 18: Learning ability 124

18.1 Theoretical background

18.1.1 The origin of the idea

18.1.2 Studies with simple learning tasks in the U.S.A.

Page

18.1.3 Studies with complex learning tasks in the U.S.A.

18.1.4 Studies with complex learning tasks in Germany

18.1.5 Intelligence and learning ability

18.1.6 Personality variables and learning ability

18.1.7 Our own studies with programmed learning material in Hungary

18.2 The Budapest study

18.2.1 Method

18.2.2 Treatment of data

18.2.3 Expectation

18.2.4 Results

18.3 The Sherbrooke study

18.3.1 Scoring

18.3.2 Treatment of data

18.3.3 Results

18.4 The New York study

18.4.1 Method

18.4.2 Treatment of data

18.4.3 Question of interest

18.4.4 Results

Chapter 19: Creativity 172

19.1 Theoretical background

19.1.1 Creativity and intelligence

19.1.2 Personality correlates of creativity

19.1.3 Fostering creativity in the school

19.2 The Budapest study

Page

19.2.1 Method

19.2.2 Treatment of data

19.2.3 Expectation

19.2.4 Results

19.3 The Sherbrooke study

19.3.1 Method

19.3.2 Treatment of data

19.3.3 Results

Chapter 20: Perseveration and Sociometric structure 224

20.1 Perseveration

20.1.1 Theoretical background

20.1.2 The Budapest study

20.1.3 The Sherbrooke study

20.2 Sociometric structure

20.2.1 Theoretical background

20.2.2 Method

20.2.3 Treatment of data

20.2.4 Expectation

20.2.5 Results

Chapter 21: Attitude 237

21.1 Theoretical background

21.2 Method

21.3 The Budapest study

21.3.1 Method

21.3.2 Treatment of data

Page

21.3.3 Expectation
21.3.4 Results
21.4 The Sherbrooke study
21.4.1 Method
21.4.2 Treatment of data
21.4.3 Expectation
21.4.4 Results
21.5 The New York study
21.5.1 Method
21.5.2 Treatment of data
21.5.3 Results

Chapter 22: Anxiety - Achievement motivation - Effects on personality 273
22.1 Anxiety
22.1.1 Theoretical background
22.1.2 The Budapest study
22.1.3 The Sherbrooke study
22.2 Achievement motivation
22.2.1 Theoretical background
22.2.2 Method
22.3 Effects on personality
22.3.1 Theoretical background
22.3.2 Method
22.3.3 Expectation
22.2.4 Results

Page

Chapter 23: Attitude toward problems 312

23.1 Theoretical background

23.1.1 Problem solving ability

23.1.2 Attitude toward problems

23.2 Method

23.2.1 The test development process

23.2.2 Description of the test

23.2.3 Instructions to subjects

23.2.4 Scoring

23.2.5 Statistical data

23.3 The Sherbrooke study

23.3.1 Method

23.3.2 Treatment of data

23.3.3 Expectation

23.3.4 Results

23.4 The New York study

23.4.1 Method

23.4.2 Treatment of data

23.4.3 Results

Chapter 24: Summary 349

Bibliography 356

Chapter 11

The Nature of Evaluation

11.1 The Origins and Latest Developments in Educational Evaluation

In his historical review of educational evaluation, J.C. Merwin (1969) points out that the problem of evaluation and attempts to deal with it first reached the scientific level in research conducted by R.W. Tyler in 1934. Although the role of evaluation in education is perhaps as old as education itself, Tyler was the first to define evaluation as a set of procedures for appraising changes in students, and proposed that educational testing concern itself with the changes in students produced by educational means.

With the advent of the space age in the 1950's, new developments in education, concerning course and curriculum content, have also brought about a need for accompanying changes in evaluation methods:

> "Since World War II, and particularly during the past decade, profound changes have been taking place in educational evaluation. New concepts, new procedures, and new instruments of evaluation are emerging from the interaction among new needs for educational evaluation, new conditions that must be met, new knowledge about education, and new technologies that can be utilized. As an example of this interaction, most theories of testing and evaluation have been developed on the assumption that the primary uses for tests and other evaluative devices are to measure individual differences and to furnish reliable estimates of the mean scores for groups of persons. The theories formulated to guide practice under this assumption have emphasized homogeneity of test content and the concentration of item difficulties near the 50-per-cent point. Since Sputnik, massive financial support has been given to projects concerned with the development of new courses in science and mathematics. Those supporting the construction of the new courses and teachers and administrators who are considering the use of them in their schools are asking for an evaluation of the effectiveness of the courses in comparison with other courses in the same fields. Most tests on the market were not constructed to furnish relative appraisals of different courses, and they have been found inadequate for the task. This need for evaluation of courses and curriculums is stimulating the development of new procedures, instruments, and theories that are designed to meet the need." (R.W. Tyler, 1969, p. 1).

Several of the more recent changes in evaluation content and concepts have been cited by Merwin (1969):

i. Status versus change

Current evaluation procedure consists of status determination with little systematic references to prior measures. However, educators who share the philosophical position that the teacher's role is to move a child from his present state of development toward his maximum potential development must agree that a measure of status at two points in time is necessary as a basis for determining change, or "growth".

ii. Breadth

As far back as the 1930's, Tyler emphasized the need to broaden the realm of evaluation to include the educational objectives to be achieved. (See B.S. Bloom, 1956; and D.R. Krathwohl, et al, 1964.) However, evaluators have only recently begun to put this idea into practice.

iii. Individual versus group

The entire concept in education of analyzing individual differences between learners, places a heavy emphasis on the individual. However, in recent years, much more attention has been given to the development of evaluation as it relates to broader considerations such as teaching materials, instruction and curriculum. This new direction can be attributed in part to the growing number of innovative curriculum programmes in several diverse educational fields. The trend is especially evident in mathematics education in such representative programmes as the School Mathematics Study Group (SMSG), the Madison School Project, and the Dienes Programme.

J.C. Flanagan (1969; p.223) has identified other trends in the evaluation effort, and cites a greater emphasis on the separate evaluation of individual students, and of objectives. Because modern educational programmes are designed to attain separate objectives for individuals of various types, future evaluation programmes will deal with how well this instruction has prepared the individual to utilize his unique talents for a fulfilling and appropriate role in society. The separate evaluation of objectives has come about because composites evaluations require an arbitrary assignment of weights for each specific objective in an educational programme, and since universal agreement of appropriate weights does not exist, an evaluation programme which provides separate measures has much greater applicability.

11.2 Evaluation Models

To better understand the evaluation model used in our study, perhaps an examination is in order of the definition and theories behind the various models in existence today.

Despite the appearance in recent years of several new books and articles on this topic (see Tyler, 1969; and M.C. Wittrock and D.E. Wiley, 1970), surprisingly little theory has been developed on which to base sound evaluation practice; and very little linkage exists between actual programme evaluations in public schools and the so-called "theory of evaluation" (see Provus, 1969). Perhaps, as R. Glaser assumes, (see Wittrock and Wiley, 1970; p.95), a practical evaluation theory does not exist at all. In fact, it can be safely stated only that evaluation practice is based on many different models and techniques. Recent literature describes several types of educational models and M.W. Wahlstrom has lately attempted to classify these efforts (see Table 11-1). Based on Wahlstrom's categorization of these model prototypes, it seems that our evaluation can be classified as a "research evaluation model", but, as we will show later, it also incorporates characteristics from the other models.

Table 11-1

PROTOTYPES OF EVALUATION MODELS*

adapted from R.E. Stake, October, 1969.

PROTOTYPE	KEY EMPHASIS	PURPOSE	KEY ACTIVITIES	KEY VIEWPOINT USED TO DELIMIT STUDY	OUTSIDE EXPERTS NEEDED	EXPECTED TEACHING STAFF INVOLVEMENT	RISKS	PAYOFF
Ralph Tyler's Evaluation Model (Tyler, 1942, 1969)	Instructional objectives	To measure student progress toward objectives	Specify objectives; measure student competence	Curriculum supervisor; teacher	Objectives specifiers; measurement specialists	Conceptualize objectives; give tests	Oversimplify school aims; ignore processes	Ascertain student progress
School Accreditation Model	Staff self-study	To review content and procedures of instruction	Discuss program; make professional judgements	Classroom teacher; administrator	None, unless authentication by outside peers needed	Committee discussions	Exhaust staff; ignore values of outsiders	Increase staff leadership responsibility
Bob Stake's (1967) Countenance Model	Description and judgement data	To report the ways different people see the curriculum	Discover what audience wants to know about; observe; gather opinions	Audience of final report	Journalists; social psychologists	Keep logs; give opinions	Stir up value conflicts; ignore causes	Broad picture of curriculum and conflicting expectations
Dan Stufflebeam's CIPP Model (1971)	Decision-making	To facilitate rational and continuing decision-making	Identify upcoming alternatives; study implications; set up quality-control	Administrator; director	Operations analysis	Anticipate decisions, contingencies	Overvalue efficiency; undervalue student aims	Curriculum sensitive to feedback
Research Evaluation Model	Cause and effect relationships	To seek simple but enducing explanation of what works	Exercise experimental control and systematic variation	Theorist; researcher	Research designer; statistical analysis	Tolerate experimental constraints	Artificiality; ignore personal values	Get rules for developing new programs
Formative Evaluation	Timely and continuous feedback	To facilitate improvement of programs over short time intervals	Continuous monitoring and reporting of ongoing activities	Program Developer	Curriculum Specialist	Implementation and evaluation of program	Costly	Program changes incorporated from evaluation activities at relevant times
Summative Evaluation	Outcomes and costs	To obtain terminal decisions about programs	Defining and measuring outcomes and costs	Program Director of Administrator	Measurement Specialists; Sociologists; Economists	Facilitate data gathering and program implementation	Over-emphasing on ends at expense of means. No inputs for program changes	Composite decision at conclusion of program

11.2.1 Definition and Classification

G.T. Evans (1969) has determined that, in general, tests and examinations in the educational system are designed to serve the following ten purposes:*

> 1./ *Feedback:* to assist the student and his teacher to know how he is progressing toward some desired goal
> 2./ *Diagnosis:* to indicate areas of relative weakness and, if possible, to indicate ways of correcting such weakness
> 3./ *Self-knowledge:* to help the student understand what kind of person he is, what he does well, and what he does badly. To allow him to arrive at reasonable and realistic levels of aspiration in his vocational and educational choices
> 4./ *Readiness:* to assess the student's chances of success in a new course of study or in a more advanced unit of work
> 5./ *Vocational and educational guidance:* to help the student choose between educational or vocational alternatives, or to indicate the appropriate next unit of work
> 6./ *Selection:* to assess the chance of success in a particular employment or particular educational institution for which there is a limited number of openings, or in which failure may be costly both to student and institution
> 7./ *Certification:* to guarantee that the student has reached a minimum level of skill in areas in which he has vocational or public responsibility
> 8./ *Teacher feedback:* to help teachers understand the strengths and weaknesses of their teaching programs
> 9./ *Course and curriculum evaluation:* to help judge the efficiency of a particular curriculum in achieving its goals
> 10./ *Educational research:* to provide the quantitative variables with which to explore the educational process
>
> The first five of these purposes are clearly pupil-centered. They satisfy needs that seem to be essential to the process of deliberate and informed educational treatment of any particular child. They seem as necessary to the teacher as being able to see is to an artist, and important to the student's self-knowledge and control over his development.

The last five are less concerned with the individual pupil than with the teaching task as a whole, and with the needs of society. Therefore, viewed in this framework, our evaluation study can be classified under the ninth category: course and curriculum evaluation.

11.2.2 Judgements versus decisions.

In further defining the purposes of evaluation, Wittrock (1970; p.3), differentiates between evaluation for *making judgements* and evaluations for *making decisions* concerning instruction or the content of instructional programmes. Another viewpoint defines evaluation efforts as the *assessment of pupil achievement* and the use of this information for making decisions about individual students or for making decisions about the quality of instruction in general. Another definition is the gathering of many types of pupil data, and of which achievement measures are only one aspect used for the purpose of planning learning activities suited to the individual. Evaluation also describes the study of teacher performance for the purpose of improving it(C.M. Lindwall and R.C. Cox, 1969, p. 156). In a more specific context, E.G. Beale and J.W. Wilson see educational evaluation as the quality of the organization of the *mathematics material* presented to the pupil and to study whether this material needs changes. Reserving this last mentioned kind of evaluation for mathematicians and mathematic educators, our study deals directly with *what identifiable changes in students can be observed after exposure to the mathematics programme*?

11.2.3 Summative and Formative Evaluations.

It is quite common nowadays to discriminate between "*summative*" and "*formative*" evaluation studies (C.F.M. Scriven, 1967(b)), where the former applies to a final evaluation of a new approach with a conventional one, and the latter refers to evaluations carried out during the process of development of a new programme (e.g. Flanagan, 1969). Because the Dienes programme has evolved into a well developed system, but nevertheless at the same time is in a continuous state of flux, our study cannot be placed in one or the other of the above described categories as defined. However, following the basics laid down in Bloom's 1968 article "Learning for mastery", J. F. Weaver presents these two concepts in a different light: *summative evaluation* as a separate process from teaching and learning, while *formative evaluation* as an integral part of instruction (see Fig. 11-1). In this respect, our evaluation study can be categorized as summative although some of the measures used (e.g. the Secret Writing Test: the Color/Shape Secret; and the Sherbrooke Problem Facing Questionnaire; see later) can accurately be considered as part of the curriculum.

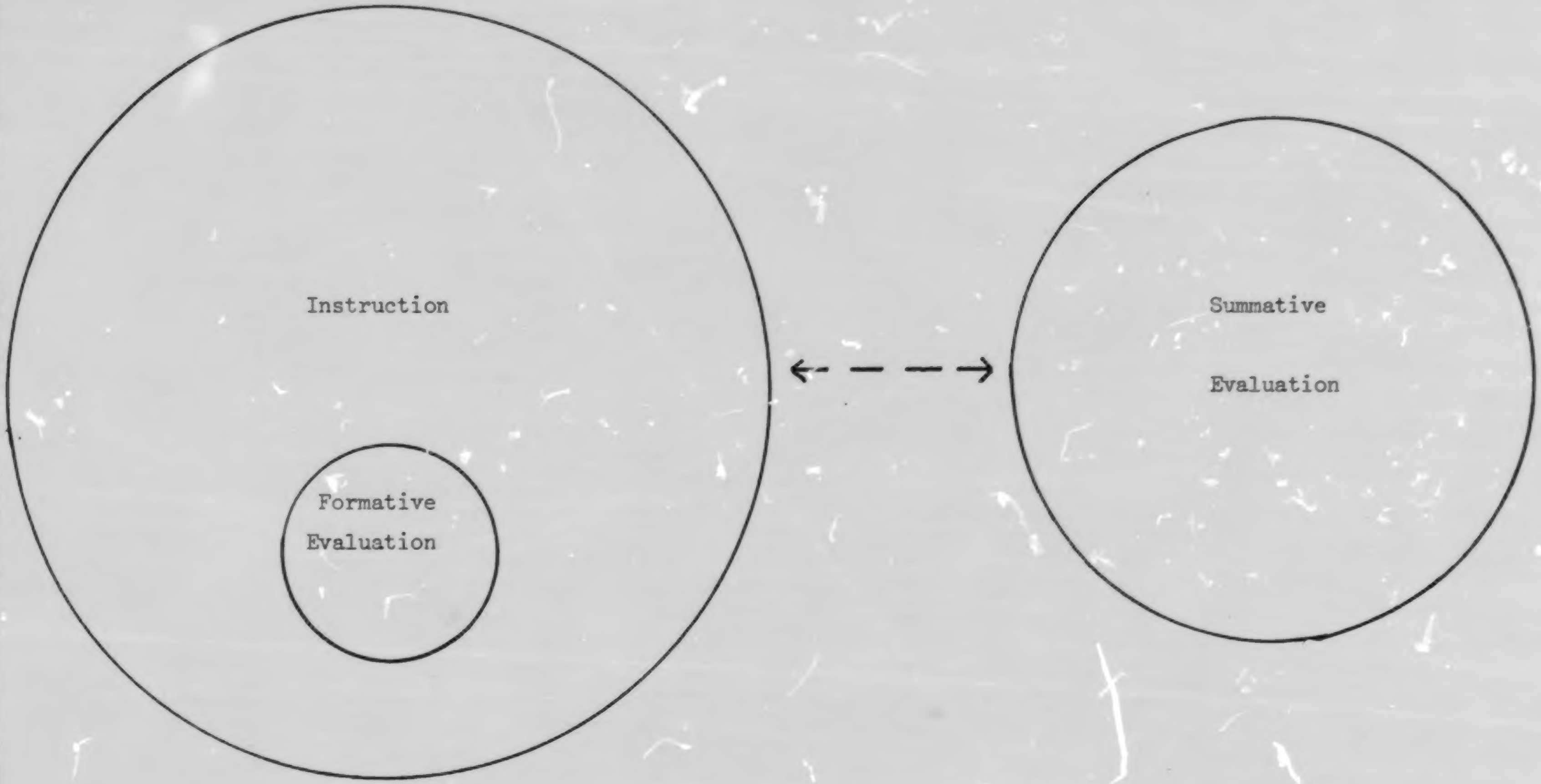

Instruction
Formative
Evaluation
Summative
Evaluation

Fig. 11-1 Instruction, formative and summative evaluation

11.2.4 Informal evaluation.

Wittrock (1970) mentions a further distinction found in conceptual approaches to evaluation: the explicitness of the variables and the behaviour they are concerned with. For example, in an informal evaluation situation, judgements expressed are not necessarily backed by explicit statements explaining the basis, values, experiences, variables, and data on which they are based. After several visits to schools using the Dienes method of mathematics learning, our researchers were "excited" with the classroom environment in which children were free to work alone or in groups, solving several different problems or creating new tasks. Initial reactions to this environment often brought out comments such as "great" or "much better than traditional arithmetic learning techniques". However, in comparison it is certainly understandable that when someone else, perhaps a parent of one of the students enter this environment , his initial reaction to the seemingly unstructured play would be to call it "terrible disorder", and may even wonder how these children, some of whom are sprawled on the floor, some arguing, some lost in private thought, are expected to acquire any degree of discipline and knowledge. This kind of judgement evaluation is based solely upon the evaluators past experiences and knowledge and therefore its value is a function of his attitudes. Of course, in the hands of a qualified person, this type of evaluation can serve as a useful tool for making educational decisions, but usually it is only a beginning step toward a more formal evaluation. In our study, this method is used only in describing the classroom environment found in the schools evaluated (see Vol. I, chapters 7, 8, 9 and 10).

11.2.5 Formal Evaluation

In a formal evaluation method, which is what we have hoped to achieve here, in addition to forming judgements and decisions, the basis on which individual judgements and decisions are founded must be explained. This need for an explicit measurement of the basis of a judgement is in fact one of the most difficult elements in a formal evaluation, requiring a great deal of skill on the part of the evaluator.

Wittrock (1970) distinguished four different types of formal evaluation:

i. The evaluation of environment, which concentrates on the physical and human characteristics of the learner's environment.

ii. The evaluation of learners, where the behaviour and characteristics of learners are explicitly outlined, usually in order to explore the differences among students.

iii. The evaluation of learning, where changes are measured in students behaviour toward behaviourally defined objectives in order to determine the student's progress during a certain period of time.

iv. The evaluation of instruction (or "cause and effect" evaluation), in which an attempt is made to relate to the variables found in the above three cases; that is, the salient characteristics of the instructional environment, the individual learners, and their actual learning.

Working within the frame of this fourth classification, Wittrock maintains that with the appropriate statistical tools, "cause and effect" relations can be estimated from the naturalistic data provided by the educational evaluation. In this respect, our evaluation study best fits this last described category.

Another helpful distinction can be made between the evaluation of procedure and the evaluation of accomplishment. While it is quite possible to evaluate a procedure, such as a difficult surgical operation, and show that it is being done properly, it is another matter completely to evaluate the beneficial results. When the evaluation of technique is neglected (as was done in our study) and the focus is moved directly to the evaluation of accomplishment, the effective implementation of the procedure is assumed (Glaser, 1970). We felt especially justified to procede in this manner in view of the fact that the evaluated experimental schools were very closely related with the theoreticians responsible for the educational innovations.

11.3 Common Elements in Modern Evaluations

From this brief survey of the many varieties of evaluation and the distinctions among them, it is possible to determine the more essential common elements in a modern approach to an evaluation effort:

i. The definition of all the outcomes of the system, including the objectives or aims and also the possible unplanned effects.

ii. The systematic analysis and study of various possible procedures for achieving the objectives as defined (i.e. a detailed description of the inputs of the system; characteristics of the students and teaching staff; facilities; instructional materials and methods; observation and measurement procedures; and available technical assistance).

iii. A plan and decision based on this analysis and an over-all evaluation of the final program. (see Flanagan, 1969; p. 222).

In order to clarify the purpose of our particular study, it is important to specify for what uses our evaluation is intended. We consider evaluation as a process enabling us to collect data about different personality factors in the students, (both cognitive and affective factors), which are most likely affected by the educational environment created to reach the specific objectives of the educational system. This data indicates the assessment of an entire curriculum and instruction system for the purpose of making decisions concerning its worth.

11.4 The Necessity for Improved Evaluation Methods.

11.4.1 The Need for Innovation in Teaching Methods and Objectives.

A growing spirit of innovation is increasingly evident among those concerned with the structure and function of education at all levels and in all subject matters. Evidence of this spirit has been acknowledged in the recent publications of reports and, more significantly, in the implementation of new educational schemes and designs at the primary, secondary, and high school levels (J.B. Biggs, 1967). Today's educators realize more than ever that in order to adapt education to our rapidly changing world we need continuous and in some aspects basic changes both in curriculum and in teaching methods.

The main impetus of this innovation has been toward the modernization of teaching and learning mathematics. Although the birth of this movement toward modernization can be traced back to before the onset of the space age, it is generally believed to have surfaced fifteen years ago when the Russians launched their first Sputnik. The aftermath of this unprecedented scientific event led to the realization, especially in America, of the importance of thorough preparation for scientific personnel, specifically centering upon the necessity of improving the teaching of mathematics beginning at the primary school level (T. Varga, 1965). As Z.P. Dienes and E.W. Golding (1967 p.1) emphatically state, "We need to change from our traditional courses in mathematics because it is already obvious that these cannot serve the needs of the world of tomorrow, as they did the world of yesterday."

11.4.2 The Need for Feedback Through Evaluation.

A "change agent" who introduces a "planned change" - that is, a change which is derived from a purposeful decision to effect improvements in a personality system or a social system and which is achieved with the aid of professional guidance (R. Lippitt et al., 1958) - must receive feedback concerning results of such a change. There is no reason why educators should be an exception to this rule. (J.D. Williams, 1971). The growing number of new programmes in practically all subject matters demand in the way of feedback more studies for the measurement of the effects of different teaching materials, instruction, and curriculum (e.g. in geography, see N. Senathirajah and J. Weiss, 1971; in medical education, see P. Kendall, 1964). This need is especially evident in mathematics learning (see e.g. Merwin, 1969, Reys and Post, 1973). Because of the radical changes in the curriculum at all grade levels, coupled with a variety of new pedagogical techniques, the evaluation of mathematics programmes is considerably more important today than was the case a decade ago (Begle and Wilson, 1970). In much the same vein, R.E. Stake questioned this lack of information as early as 1967: "One of the biggest investments being made in the U.S. education today is in the development of new programmes. School officials cannot yet revise a curriculum on rational grounds, and the needed evaluation is not under way. What is to be gained from the enormous effort of the innovators of the 1960's if in the 1970's there are no evaluation records".

Without an evaluation process, the innovator, even in the most capable case, can express only his own feelings; for example, as W.P. Hull (1963) does: "We can only say that, intuitively, we feel that we are on the right track," or as R.B. Davis writes: "We do not know whether we are on the right track or not...". Obviously, personal conviction is in certain cases not enough. In most countries, before any new method of education can be used (at least to any great extent), there must be some "objective" proof of its advantages. Stake (1967) emphasizes this in his explanation of M. Scriven's (1967 b) opinion (The evaluator is the one who should judge education systems.):

"Different decisions must be made - P.S.S.C. or Harvard Physics? - and they should not be made on trivial criteria; e.g. mere precedent, mentioned in popular press, salesman personality, administrative convenience, or pedagogical myth." Therefore it is imperative that we face the problem of evaluation in order to successfully confront the problem of improving education.

Chapter 12

Difficulties in educational evaluations

While the necessity for comprehensive evaluations is greatly increasing, it is also true that those educators who wish to carry out evaluations are very often confronted with several obstacles (see J.D. Williams, 1963; 1965; 1967; 1971). For example, teachers' abilities (their experience with the method used, their enthusiasm or otherwise for the method, etc.) are liable to influence results considerably. And the effect of experimental stimulation is difficult to control, especially in short term investigations. Matching pupils in experimental and control groups in terms of mathematics performance is a particular problem. And if the new system is a complex one (changing the method and the curriculum together - and there is no other way of educational improvement possible), then it is difficult to distinguish the factor which causes the result. Unplanned changes can take place in the experimental situation, and these must be taken into account. The instruments of evaluation can turn out to be either inappropriate or without known application. Throughout the evaluation process it is possible to ignore the special merits of the educational project, or, on the other hand, to discover trivialities about it. The more noteworthy student achivements are unlikely events and are easily masked by statistical noise in testing procedures. In education - as in quantum physics and brain surgery - the act of measuring causes a significant alteration of the situation.

Evaluation is primarily concerned with changes in individuals (or as in most cases in groups of individuals) (B.S. Bloom, 1970). The measurement of change presents numerous unsolved problems (see e.g.

C.W. Harris, 1963), especially in scaling and reliability of testing. For example, the reliability of gain scores are mostly sufficient for research conducted on groups but rarely for the purpose of individual measurement (J.P. Guilford, 1970). In most school operations, the experimental materials are often used by only the better teachers with above average students and are not adopted on a wide scale until an explicit evaluation design is accepted and followed by the school (Begle and Wilson, 1970).

The inability to solve these and other difficult problems is partly responsible for the present tendency to spend even less effort on evaluation than some years ago. Problems like these are real and often unsolvable.

While some psychologists have lost interest because they cannot control the variables closely enough many "educationists" have become dissatisfied with evaluation processes because of their "sterility". There is certainly a tendency in psychological experimentation toward increasingly trivial meaningless investigations which are irrelevant to human problems (see R.N. Sanford, 1965). As Varga (1968) summarizes: "Paradoxically, what would be really worth evaluating - the growth of understanding original creative thinking - substantially escapes the present-evaluating procedures. Any sufficiently exact evaluation is restricted to formal skills, solutions to pre-determined categories of problems, retention, etc. Consequently, too much stress on evaluation entails a shift towards less important aspects of the work."

Because education is a complex process, resulting in equally complex and often subtle outcomes, it is not unusual for writers to cite the existence and importance of these outcomes, often suggesting that the intangible outcomes of education are difficult or even impossible to measure (see R.L. Ebel, 1963). Contrary to this common opinion, our research has led us to believe that all important outcomes of education are potentially measurable. (If attaining an alleged goal of education does not change the overt behaviour of the person who attains it in any way, on what basis can it be said to be important?) Of course, this certainly does not mean that all the important outcomes of education can be measured easily. As Ebel points out:

"The practical limitations of effective educational measurement are real and many, as anyone who had labored to improve educational measurements can testify. But they are not fixed eternally by the nature of man, nor is it useless to try to overcome them. The possibility of measuring the degree of attainment of all important outcomes of education does exist."

The rest of this chapter deals in detail with two of these difficulties. In doing so, we affirm our conviction that the time is near when educators and psychologists will be able to overcome these difficulties. We have to in order to avoid the educator's predicament described here:

"You can't measure the effects of what I do.

-- Why not?

They're intangible.

-- Oh? Why should I pay for intangible results?

Because I've been trained and licensed to practice.

-- Hmm... all right. Here's your money.

Where? I don't see it.

-- Of course not,... it's intangible." (R.F. Mager, 1962).

12.1 Finding Behavioural Objectives.

> Once upon a time a Sea Horse gathered up his seven pieces of eight and cantered out to find his fortune. Before he had traveled very far he met an Eel, who said,
>
> "Psst. Hey, bud. Where 'ya goin' ?"
>
> "I'm going out to find my fortune," replied the Sea Horse, proudly.
>
> "You're in luck," said the Eel. "For four pieces of eight you can have this speedy flipper, and then you'll be able to get there a lot faster."
>
> "Gee, that's swell," said the Sea Horse, and paid the money and put on the flipper and slithered off at twice the speed. Soon he came upon a Sponge, who said,
>
> "Psst. Hey, bud. Where 'ya goin' ?"
>
> "I'm going out to find my fortune," replied the Sea Horse.
>
> "You're in luck," said the Sponge. "For a small fee I will let you have this jet-propelled scooter so that you will be able to travel a lot faster."
>
> So the Sea Horse bought the scooter with his remaining money and went zooming thru the sea five times as fast. Soon he came upon a Shark, who said,
>
> "Psst. Hey, bud. Where 'ya goin' ?"
>
> "I'm going out to find my fortune," replied the Sea Horse.
>
> "You're in luck. If you'll take this short cut," said the Shark, pointing to his open mouth, "you'll save yourself a lot of time."
>
> "Gee, thanks," said the Sea Horse, and zoomed off into the interior of the Shark, there to be devoured.
>
> The moral of this fable is that if you're not sure where you're going, you're liable to end up someplace else.*

Before we can effectively measure the outcomes of education research, we must define the objectives of the research in an initial planning stage.

> Obviously, the evaluation of an experimental program presupposes fairly precise knowledge of the objectives of the program. Without such knowledge, it is manifestly impossible to say whether or not the program has been successful. (P. Kendall, 1964).

---

* From the preface of R.F. Mager's book (1962), Preparing Instructional Objectives.

Because any course in instruction is intended to help students in the acquisition of certain skills and abilities, it is absolutely essential that an attempt be made in the planning stages to analyze and spell out objectives as well as the learning experiences needed to bring them about (R.Wood, 1968). As with all enterprises which seek to improve upon present practice, this is a difficult task to bring off, but it is completely necessary for any innovation or research project in education. Activities which require the time and efforts of both students and teachers are justifiable only if they contribute to the attainment of educational goals. In this respect, most professional educators would agree that not only evaluation, but every aspect and arrangement of the educational enterprise, including facilities, staff, materials, organization, administration, curriculum, instruction and guidance, should be selected and operated to contribute to the attainment of educational goals (Ebel, 1963). Since the primary purpose of the evaluation is to indicate the success or failure of a certain school programme, it is especially important to put the major emphasis on the educational objectives.

In the 1930's, Tyler advocated a close tie between objectives which serve as a basis both for instruction and evaluation, and proposed some rather specific guideline for the formulation of these objectives. One of his guidelines suggests that objectives should be spelled out in statements of student behaviour which should then serve as the objectives of teaching and the basis for testing. He further proposes that our very limited view of evaluation of specific elements of knowledge should be broadened to encompass all important outcomes of education. The techniques employed for evaluating "higher mental processes" are described in detail. In addition to the cognitive outcomes, including knowledge and higher mental processes, we should reach for and evaluate affective outcomes of education. (Merwin, 1969).

The main problem confronting evaluators who wish to define the objectives is that schools are generally not articulate or clear enough concerning their desired instructional aims, even concerning those innovations in which the school takes particular pride (H.M. Brickell, 1969). It is true that specified objectives can have a constraining effect on education (E.W. Eisner, 1967) while implicit objectives can direct an educational programme into the most productive direction; however, it is extremely difficult to communicate implicit objectives to other people (parents, teachers in other schools, etc.) Evaluators constructing instruments for use by more than one teacher or one school also find it difficult to start work until a set of specifications are prepared. If these specifications are not available, evaluators must themselves construct them or create a committee or another type of consensus mechanism to develop a set of useful specifications against which they can construct and validate an evaluation instrument. The clearer the specifications are in terms of behaviour and content, the better, because they are most useful in defining the problems to be solved in the construction of evaluation instruments, and also because they provide the criteria against which the instruments are validated (B.S. Bloom, 1969). Why the statement of these specifications is so important is discussed in length by R.E. Traub (1969):

> From the point of view of an evaluator, it is desirable to state objectives as specifically as possible in terms of behavioral change:
>
> (1) Specifying change is important because education becomes redundant if the desired behavior exists prior to commencement of teaching-learning activities. (2) Behavioral objectives tell the evaluator what behavior to look for as evidence of achievement, and guide the designer of situations for observing attainments. For example, the goal of teaching pupils to add up to five two-digit positive integers in which "carrying" is required, given that they know how to add five two-digit positive integers in which no "carrying" is required, is a statement that suggests very clearly the kind of addition problems that must be used to ascertain whether the objective has been attained. (3) Properly stated objectives also imply a standard of performance. In the example just cited, a certain set of questions is specified, any of which the student would be expected to answer correctly after completion of appropriate teaching-learning activities. Of course, perfect performance of the total set of problems may be viewed as too stringent a criterion, in which case the operational statement of the objective could specify a more moderate goal, say successful completion of 80% of all the problems presented. (4) Specific operational goals guide to some extent

> the design of teaching-learning activities. At the very least, they provide a preliminary basis for judging whether the behaviors expected to result from the teaching-learning activities under consideration bear any similarity to the objectives being pursued.
>
> The attempt to define educational objectives in operational terms is itself likely to turn up interesting results. An objective, or some aspect of it, may not be amenable to analysis and interpretation in terms of behavior, and for this reason may be dropped from further consideration. Alternatively, the process of operationalizing an objective may result in the identification of another objective not originally listed, but which seems nevertheless to be important. In this way, the list of objectives will become refined .*

We feel that objectives stated in this manner, as they are in the Dienes mathematics programmes, should be made more explicit and then expressed in behavioural terms before the degree of their achievement can be assessed (see e.g. Mager, 1962) Explicit statements from the innovator-evaluator concerning priorities assigned to measures, and relating facts to which evaluative dimensions are included in the study are crucial (L. Cahen, 1970).

12.1.1 Aims

In recent literature (see e.g. Wood, 1968), there is a tendency to make a sharp distinction between aims and objectives. Aims are general declarations of intent which, if needed, can give both direction and shape to a teaching programme but are usually insufficiently precise to be of much value for curriculum development or evaluation since they do not specify behaviours which may be subsequent foci for the development of appropriate classroom procedures (for example ; "Schools should prepare our children for jobs in the unforeseeable future", or "Students should be equipped with mathematics which enables our country to compete in the increasingly technological 'Space Age"; etc..(see P.A. Taylor and T.O. Maguire, 1966), these authors use the terms "broad objectives" for "aims" and "interpretations" for objectives).

---

* For further discussion see e.g. P.L. Dressel and L.B. Mayhew, 1954; H.S. Dyer, 1966; D.R. Krathwohl, 1965.

These aims are usually direct offshoots of the philosophy adopted by the school system: Why do they teach what they teach?

Among the most common of the traditional justifications for mathematics education are:

i. usage in everyday life by wide segments of the population;
ii. service as a tool in other sciences ("...the natural language of quantitative and qualitative relationships, and of science and technology...");
iii. training in logical thinking ("...a concise basis for logical thinking...");
iv. appreciation of mathematics because of its beauty ("...stimulating pleasurable intellectual experience..."); and
v. part of our cultural heritage and influential in shaping our history.

These aims still hold some significance in education today. However, as the K-13 Committee for the reform of school mathematics curricula reports: "For the vast majority of the public, including a majority of the intellectual public, these values bear little relationship to their own personal experience. There is little question that the mathematically minded community is an unusally segregated one with regard to its understanding of the objectives of its own discipline." (1970; p.3) The Committee found, among others, the following new facotrs which could greatly influence our early 1970's attitudes toward desirable and realistic aims of mathematics education:

i. The economic implications of education, both to the individual and to society as a whole, are becoming more generally understood. In this respect, it is worth noting that a special investment in gifted students can be economically justified on the basis of their disporportionate economic significance to the whole community.

ii. Mathematics today plays a growing role in other disciplines. Its use is no longer reserved solely for the engineers and physical scientists, but plays an important role in natural science, social science, business, industry, and several other fields of application.

iii. A need exists for a more extensive and perhaps more suitable mathematics programme as part of the general education of the average person, in his role of potential producer as well as a citizen and consumer. Many examples can be given of the advantage taken of the general public's ignorance of logic and probability and statistics in the fields of advertising and propaganda.

iv. It would be ludicrous for us to ignore the rapid growth and development of high-speed digital computers, an essentially mathematical machine. Although there is little cause for alarm that computers will be ignored by the schools, it is still imperative, aside from practice in usage, that mathematics education provide a suitable reoriented preparatory background.

In the light of these topical considerations, the Committee found the following list of aims desirable for a universal education in mathematics:

i. The development of meaningful learning---that is, learning that enables the learner to grasp relationships inherent in the material being learned and those relating it non-arbitrarily to his own background knowledge;

ii. a) Understanding of and competence in the processes of arithmetic and elementary algebra, including the use of formulas;

b) Appreciation of algorithms and related techniques in methods of larger-scale computations;

c) Understanding of approximation and order of magnitude;

iii. Understanding of the general properties of geometrical figures and the relationships between them;

iv. Basic knowledge of graphical methods, and an appreciation of elementary statistics and probability;

v. An understanding of the deductive method as a way of thinking and the ability to recognize and apply it; this includes the ideas of assumptions, rules of inference, and methods of proof;

vi. The development of mathematical intuition;

vii. The development of problem-solving skills and strategies in finding solutions to significant and relevant problems---in other words, participation in creating mathematics;

viii. The accompanying realization that mathematics is a continuing creative endeavor with aesthetic values similar to those found in music and art;

ix. The development of sound attitudes to mathematics and its uses;

x. The opportunity to study mathematical topics to a depth limited only by ability and desire with a view to achieving a fuller understanding of mathematics and a better facility in its application.

12.1.2 Objectives

The term objectives could be used only for detailed descriptions of the kinds of behaviour which teachers and evaluators are prepared to accept as evidence of success. As J. Eggleston writes, "Aims are expressions of strategy while objectives are essentially tactical in character." Many papers have been presented expressing the aims of mathematics learning in this sense, but only a few express the objectives. (See S. Avital and J. Shettleworth, 1968; this article concerns itself with only the achievement aspect.)

The Mathematics Evaluation Materials Package Project (1972) presents an excellent example of how aims in a mathematics programme can be converted into educational objectives—at least on a limited topic: achievement in sets and numbers. According to the Project, a programme in mathematics should provide for the pupil:

i. experiences with sets and numbers, with shapes and space, and with measurement;

ii. involvement as an active participant in the learning process;

iii. learning experiences in which he can have success and satisfaction;

iv. opportunities to develop some skill in computation;

v. activities, manipulative devices, pictures, and diagrams to develop and express mathematical ideas and logical thinking;

vi. opportunities to investigate topics in greater depth as interests and abilities expand;

vii. opportunities to develop understanding of structure and patterns in mathematics;

viii. opportunities to relate mathematics to the world around him;

ix. opportunities for the development of self-evaluation;

x. situations that help to develop interest in have fun with mathematics.

The Project then presents the following related objectives, accompanied by several tasks supposingly designed to measure these objectives:

i. sorts objects into sets and gives the sorting rule;

ii. gives a verbal description of a set by distinguishing between objects that are members of the set and objects that are not members;

iii. given two equivalent sets, demonstrates a one-to-one matching between members of the sets physically associating the objects;

iv. distinguishes between pairs of sets whose members can be matched one-to-one and pairs of sets whose members cannot be matched one-to-one;

v. given two nonequivalent sets, identifies the set that has the greater number of members and the set that has fewer members.

vi. states, without counting, the number of objects in a set, up to five objects, etc.

Recently many efforts have been directed toward the development of a technique to identify objectives more precisely (Flanagan, 1969; R.B. Ashlock, 1970). These efforts proved to be especially useful and are thoroughly treated in The Taxonomy of Educational Objectives (B.S. Bloom, 1956; D.R. Krathwohl et al, 1964) which systematically describes the objectives both in the cognitive and affective domains. Teachers and educators could very well profit from such books and articles describing in detail the methods of preparing instructional objectives (see e.g. Mager, 1962; Krathwohl, 1965; R.M. Gagné, 1965). However, evaluation specialists still ignore the need to report what persons and programmes and institutions are trying to do. Stake and Denny discuss this lack:

> An evaluator's technical skill should help the educator convey his purposes, both those that quickly come to mind and those implicit in what he does. What are the present methods for getting him to formulate a statement of philosophy or a rationale; to detail the encounters he wants students to have; to specify the behavioral capabilities he wants students to attain;

> or to sketch the way the ideal individual, team, or system works? Our methods now are crude, unstandardized, and unvalidated. They should be more evocative, more sensitive than indicated by the bald request, "Please state the objectives in the following space." (1969; p. 376)

These authors emphasize that evaluators must develop a battery of standard routines for seeking out unique and subtle purposes; a conceptualization and symbolic language that will permit at least gross representation of the relative importance of the goals are badly needed. What outcome is more important and how diligently the educators work to reach these outcomes must be determined. H.H. Walbesser (1968) suggests the use of a limited number of carefully selected and well-defined verbs (e.g. identify;distinguish; construct; name; order; describe) for writing behavioural objectives. As can be seen in later chapters, we have not fully overcome these difficulties. As suggested by A.E. Traxler and R.D. North (1963; pp. 216-217), our study attempts to place the emphasis on objectives which are clearly and definitely stated (or at least we tried to make these objectives as clear as possible), and which may exert an influence on the school's day-to-day work. Our study also attempted to break down broad general objectives into specific objectives in order to serve as guides for measurement.

As in the case of most evaluation efforts, the most elaborate descriptions of objectives finally appear in the tests which schools or educational projects agree to use for the assessment of their expected outcomes. Because the objectives which are thus tested tend to be those toward which teaching is directed, the tests are well worth examining and discussing (Brickell, 1969).

12.2 Criteria and methods which should be considered in an evaluation programme

A wide variety of possible criteria exists to measure pupil outcomes. It is extremely important that mathematics programmes be evaluated over a wide range of criteria (Begle and Wilson, 1970). In addition to the achievement dimensions, the inclusion of several psychological variables into evaluation work has recently been recommended by an increasing number of educators (S. Messick, 1970). For example, Ebel (1970; p. 39), in his discussion of the relation of testing programmes to educational goals, recommends the inclusion in the school testing programmes of tests for all the educational outcomes which the school is actually working to achieve: "In so far as the school program purposefully pursues such goals (as attitudes, interests, etc.) and in so far as available evaluation techniques permit, it should include measures of non-intellectual educational outcomes in testing programs."

Bloom (1969) also states the so-called "non-specified outcomes of instruction" must be measured. These "side-effects", independent of whether they are desirable or undesirable, could indeed be the most important part of the evaluation. ("If some students develop considerable cognitive competence in a subject but learn to dislike the subject with great intensity, this is clearly an undesirable consequence and should be appraised in order to determine why it occurs and how it can be altered." (p. 39; see also Dyer, 1966)

In a symposium on the evaluation of instruction (see Wittrock and Wiley, 1970), Messick (1970) strongly urged assessment of the possible and not just intended outcomes of different instructions. In a discussion of this paper, Cahen (1970) pointed out that in reality the evaluator must face two kinds of "unintended" outcomes, 1) the so-called "side-effects", that is the outcomes which were not foreseen when the new methods of teaching were introduced, and 2) the explicitly stated unintended outcome. He derives this example from our own field; the introduction of modern mathematics curricula developed in the early 1960's. These new methods attempted to develop among other aims a better understanding of different number systems, the heuristic in problem-solving, and an understanding of mathematical algorithms, but have never intended to emphasize "traditional mathematical skills", such as accuracy in routine computations or competence in translating Roman numerals. The pioneers of modern mathematics teaching have in part exposed this in their value systems.

One of the problems which came to light out of this second type of unintended outcomes was that the developers of new curricula felt—and with good reason—that standardized testing instruments have failed to measure the dimensions in which they were most interested but did in fact measure the "unintended" outcomes with relative precision and validity. This problem led to considerable debate concerning the role of comparing outcomes across competing curricula or instructional systems when the competing systems have different intended outcomes, and no real solution has been achieved until quite recently.
We occur with Cahen that summative evaluation (which is designed to provide information to the potential consumers of the instructional product) must include both kinds of dimensions, especially because we view instruction as an ongoing activity and consequently summative evaluation as a first stage of a formative evaluation for the second wave of innovation (M. Scriven, 1967b).

If an evaluation study concentrates on only one dimension, it is impossible to assign an overall ranking of worth for the instructional systems evaluated. (See e.g. the very interesting study by W.A. Brownell, 1967; 1968, with third grade pupils, evaluating the Cuisenaire, Dienes and conventional ways of mathematics teaching on the dimension of maturity level of underlying thought processes and the amount of understanding that has been attained.) Perhaps typical of today's educational innovator's attitude toward evaluation is the reasoning in the following proposal set up to establish a secondary evaluation research project in New York's Manhatten (Committee for a Comprehensive Education Center, 1970):

> Evaluation will be as difficult as it is important. We need to be careful that, out of a desire for rigorous data, we do not place harmful strictures on the experiment. We see, for instance, limited significance in standardized tests, whether the City's, the Regents', or others. If such tests are made a major part of the evaluation, that inevitably will distort the program. Evaluation, if it is to be really searching and instructive, must ask many kinds of questions.

(Of course, we do not mean to imply that useful information can be gained only from this kind of evaluation; see S.S. Willoughby, 1968.) Therefore, an evaluation which enables decisions concerning different projects should encompass all the important variables to make possible such a ranking.

However, the difficulties of administering non-achievement batteries in the evaluation of educational programmes should also be considered. Techniques such as measures of personality, cognitive style, attitude, etc. are sometimes considered to be an invasion of privacy (Cahen, 1970), by both students and school administrators, who often are not able to grasp the relevance of non-achievement type tests to the evaluation of instructional outcomes, etc. We agree largely with T. Varga that "since the evaluation of the more valuable aspects of mathematical learning is limited, the weight of these aspects in evaluation should be increased." (A. Servais and T. Varga, 1971; p. 30) Thus, our main interest was not in the mathematics knowledge the students gained throughout their 2-5 years, but more particularly in various conceptual and personality factors. The inclusion of the measurement of side-effects presents two new problems, whose origins are found paradoxically in the super-abundance and in the shortage of the methods and dimensions available.

12.2.1 Super-abundance in evaluating dimensions and methods

It has become apparent to the evaluator that there is an almost infininite number of possible dimensions to assess and evaluate. The innovator-evaluator must decide which dimensions have the greatest potential for providing information for himself as well as providing multi-dimensional outcome measures for the potential consumer (Cahen, 1970). A vast number of methods are available to measure several of these dimensions. Since 1938, succeeding editions of the Mental Measurement Yearbook have been a dependable source of expert opinion to guide consumers in the selection of tests, containing information on 2,800 tests. In the short period between 1954 and 1964, over a thousand tests of mental measurement were published (O. Buros, 1970). R.B. Cattell and F.W. Warburton (1967) is another valuable source, listing 2,366 objective tests in psychology and motivation only. It is extremely difficult to choose the appropriate dimensions and criteria out of the many possibilities, primarily because schools usually are not ariticulate enough in expressing their instructional objectives, even concerning those innovations in which they take particular pride (see Brickell, 1969), and also because of the lack of an independent fact-finding and test evaluation agency (M.D. Engelhart and J.M. Beck, 1963).

12.2.2 Shortage of relevant measurements

In spite of the existence of plentiful methods, it is an extremely difficult task to find measurement tools based upon their relevance to the educational objectives rather than upon their mere availability (which usually only dictates the choice of tools, because of a lack of time, resources, or even skills; J.K. Hemphill, 1969). This shortage of relevant measurements may well be the reason why those few evaluators, who have not concentrated solely on students' achievement, have emphasized cognitive behaviour (behaviour associated with "thinking") even in the case of mathematics instruction, where affective behaviour is generally believed to be important (see e.g. J.F. Weaver, 1970; but with some notable exceptions, e.g. Biggs, 1962).*

---

* Practically no emphasis has been given to the evaluation of the psychomotor behaviours (that is those associated with physical action), beacuse of a general belief that they are not significantly encountered in mathematics instruction. However, Weaver (1970; p. 336)

The problem of shortage in relevant measurement is described by Williams (1967) as the dilemma of instruments of evaluation: inappropriate or without known implication? Home-made tests, designed as directly relevant to the objectives of the school programmes evaluated, often prove more useful, providing a much clearer result, than "borrowed tests" of questionnable relevance (Hemphill, 1969). However, "it is no more sufficient just to construct a new test than it is just to invent a new word. In much the same as a new word gains meaning with every use, so does a test. Until a test has been used in a variety of contexts, and has thus gained some 'experience', it may be difficult to say what implications should be read into the scores it yields." (c.f. Williams, 1967; p. 190)

.....contd.)

describes a recent approach to geometry through construction, intended for children in the primary grades (I and II). The approach did not meet with much of a success or wide acceptance, due largely to the inability of many six- and seven-year olds to manipulate compasses for the drawing of arcs and circles. In mathematics instruction, such as employed in the Dienes method, where the process of abstraction and generalization begins with the manipulation of and playing with objects (S. Klein, 1972b), the study of children's psychomotor behaviour is quite relevant, and recently, J.D. Williams presented a "research system" at the 1972 annual meeting of the International Study Group for Mathematics Learning in Florence, proposing a study of the relationship between concept formation and the design of learning environments, especially the apparatus used in the Dienes method (Dienes blocks. logic blocks, etc.).

The lack of appropriate instruments has led us to develop new ones, because of the enormous work necessary to invent a "meaningful" test, and because of the limited time and research sources available for evaluation projects. However, we still are not able to solve the very practical problems of adopting enough existing tests, or even developing or validating as many new measurement procedures as thought necessary. (Recently, Warburton, 1970,gave a very expressive description of the difficulties in test-construction in connection with the construction of the British Intelligence Scale.)

For example, we were very much impressed with Messick's 1970 study of the importance of assessing possible and not just intended outcomes. Nine examples of the dimensions in *cognitive style* are given: a person's typical modes of perceiving, remembering, thinking, and problem-solving as important criteria variables in the evaluation of instruction . A very realistic possibility exists which shows that the Dienes method, after a long period of use, has a definite effect on children's cognitive style; e.g. changing it in the direction of field independence (experience items as discrete from their backgrounds and reflects ability to overcome the influence of an embedding context; H.A. Witkin et al, 1962); cognitive complexity (tendancy to construe the environment in a multidimensional and discriminating way; Bieri, 1961; Day, 1967b,c; 1968); tolerance for incongruous or unrealistic experiences (willingness to accept perceptions at variance with conventional experience; G.S. Klein et al, 1962); etc.

Many interesting studies have been conducted (see e.g. L. Muller-Wills, 1967; T.G. Decarie, 1965; J.H. Flavell, 1963; J. Hunt, 1961; Hunt et al, 1967; M. Laurendeau and A. Pinard, 1962; E.V. Sullivan, 1967; R.D. Tuddenham, 1970) on the effects of different learning experiences during the time children spend in the Piagetian stages. We feel that the Dienes method has an effect on this aspect as well.

Another important variable to be considered is curiosity. Recent studies show that intelligence and curiosity are two necessary conditions for creativity (Day and Langevin, 1969). Some attempt to measure children's curiosity have also been tried (see e.g. Day, 1967a; 1968; Penney and McCann, 1964; for some of the difficulty in the measurement of curiosity see e.g. Langevin, 1970; Wilson et al. 1972). Because of the many different new materials used by the children, it is likely to expect a difference between children in experimental and control classes with this variable. In fact, we hope that in the near future we will have the opportunity to carry out a study focusing primarily on these fields. However, an undertaking of this nature exceeds our present manpower and time limit situation.

The problems of inventing, developing, and validating measurement procedures and instruments present a large area of concern shared by researchers interested in child development and by evaluators of educational programs. For this reason, Ashlock and Herman (1970) in their collection of papers on current research in elementary school mathematics, included articles on test development in the evaluation section (e.g. one such study by F. Flournoy, 1968, on the development of an Arithmetic Understanding Test).

The connection between research and evaluation is a complicated one (see the discussion on this topic in Wittrock and Wiley, 1970; p. 27). On one hand, there is always "some concern in evaluation studies about the source of obtained effects, as well as their size, and about the generalizability of the findings---to other types of students, other time periods, and other settings, if not to other materials embodying similar principles. In this view, indeed, evaluation is research." (Messick, in Wittrock and Wiley, 1970; p. 217). In spite of this and the fact that precision, care, discipline, and logical thought are marks of a "good" evaluation as well as of "good" research, evaluation studies cannot be viewed simply as a subset or a special type of research study, and should not be judged on the same basis as that on which research studies are now conventionally rated. To do so would result in a situation in which evaluation studies had a reputation of being poorly conceived and executed research studies (Hemphill, 1969; p. 218). Hemphill further suggests that the statistical decision theory provides the best framework for evaluation studies. If we accept this, it follows that the criterion for the value of an evaluation study is to be found in its contribution to a rational decision process in situations in which it is necessary to estimate the probability of a desirable but uncertain outcome of an action chosen from a number of alternative actions.

Chapter 13

Theoretical model of the evaluation for the Dienes programmes

The theoretical view of the Dienes programmes taken by the evaluation project is outlined in Table* 13-1. As seen by Hunt (1970b), it could be divided into four fields:

i. the objectives of the Dienes programmes;

ii. the educational environment created by the Dienes programmes to reach these objectives;

iii. the conception of the person adopted by the Dienes programmes; and

iv. the evaluation programme's criterion for change in the personality.

Chapters 7, 8, 9 and 10 in Volume I discuss in detail the objectives and the environment provided with the Dienes programmes evaluated and the criteria used in the evaluations. In this chapter, we wish to discuss briefly only the problems of objectives and personality concepts.

13.1 Objectives

To verify whether an educational project is successful or not, it is necessary to have some well-defined criteria against which success may be judged. The question immediately arises whether the criteria used are to be the objectives defined by the project itself, or some externally imposed norms, perhaps supplied by the evaluator or found in the surrounding society. This question is not so clear as it may seem, as we shall see that objectives, when they are defined at all, tend to be shared by all concerned.

In the particular case of the Dienes programmes not all objectives are explicit, especially before the beginning of the evaluation process. But, as Dressel wrote in 1967, "unstated objectives implicit in practices often become far more important as guides to learning than do...explicit objectives." The first work of the evaluation process, then, was to draw out the objectives disguised in the various writings and practices of the Project.

---

* This Table will be found at the end of this chapter.

The objectives of the traditional system would seem to have been ignored in the studies. However, projects, no matter how far removed they may be from the traditional system, nevertheless grew out of it and could be interpreted as responding to needs perceived within the traditional system but largely unfulfilled by it. For example, in an unpublished report describing the Hungarian project, Varga (1970) stresses the identity of the objectives with those of the Ministry of Education's Reformed Plan of Tuition. It could be argued that in fact modern conditions have made necessary the encouragement of creativity, the reduction of anxiety, etc. and that any educational system which does not work in these directions has failed.

At any rate, the objectives in terms of which we have compared the success of the two systems of mathematics teaching are listed in Table 13-1;for the most part, they have been taken by the evaluators from the projects' priorities. The evaluations therefore show whether the different Dienes projects fulfill their objectives better (or worse) than a traditional method fulfills the same objectives, or as in the New York study , which Dienes project works better.

As Dressel writes in his study (1967), "Those who view the educational process as an extremely complicated phenomenon and regard objectives as hypotheses regarding the significant changes which it is possible to induce in individuals through the educational process, might readily admit that some objectives may be unrealistic, idealistic, or at least incapable of attainment under the circumstances in which the formal educational programme operates." This was just one of the reasons behind our evaluation programme: We wanted to see whether the Dienes projects did or did not reach their objectives.

13.2 Personality concept

The research worker's conception of personality is likely to pervade and influence an appreciable part of his investigation. The present one is subject to limitations of the researcher's conception of which personality factors are important enough to be studied, which are likely to be affected by the experimental conditions, and which are capable of being measured. As we were dealing with the effects on personality of a learning situation, we chose from the various personality concepts (see e.g. Levy, 1970) a kind of learning-theory conception of personality like Dollard and Miller's (1950). Moreover, the mere use of tests of separate personality factors on a classroom basis could lead to the view of a class as a collection of personality factors distributed in some fashion among its members, rather than as the meeting and interaction of individuals, which it also is. Again, this may be a useless theoretical distinction; in fact, a class is both of these and more; it is our semantics and our measuring instruments which force us to divide what is in fact whole.

Whether it takes place in terms of a class or of a single person, the separating off and measuring of personality factors gives merely a sample of the whole. So this experiment necessarily leaves open the question of the effects of the special learning situation on other equally interesting personality factors. (For further details, see chapter 22.) The scope of the present evaluation is limited but we hope for some more interesting results from further studies covering more factors of both cognitive (ability to abstract, ability to generalize, take short cuts, switch from irrelevant method habits, adopt new symbolism, tolerate complexity, etc.) and orectic (self expression, sociability, etc.) spheres of the personality.

As is apparent from these considerations, differentiation among the objectives as well as differentiation among the environmental factors created to realize these objectives is somewhat artificial; objectives and the means to their attainment are much more interdependent than separate, so we could expect strong correlations among the personality factors we measured. For example, there are significant correlations between anxiety and rigidity (see Kitano, 1960) and between anxiety and attitude toward school, especially amongst highly intelligent children (see Phillips et al 1960); there is an interesting relationship between perseverative behaviour, anxiety, and rigidity (Wesley, 1953); there is a stonger relationship of attitude to performance in the case of arithmetic than in that of any other school subjects (see Biggs, 1962; etc.).

13.3 Hypotheses

In accordance with the objectives given in Table 13-1, the following hypotheses were examined in the Budapest and/or Sherbrooke Study:

1. Although the children in the experimental classes spend much less time on the traditional topics, their knowledge, even in this field, is not worse than that of their schoolmates.

2. The children in the pilot classes gain a great deal of knowledge in the more and more important field of modern mathematics.

3. The children in the pilot classes are not more tired after their mathematics lesson than the children learning arithmetic in the traditional way.

4. The mathematics lesson in the Dienes classes causes a greater concentration of the children's attention than those in the control classes.

5. The Dienes method of mathematics learning helps to develop the students' learning ability more than the traditional methods.

6. The Dienes method of mathematics learning causes greater student creativity than the traditional methods.

7. The Dienes method of mathematics learning results in less perseveration in the student than the traditional methods.

8. The Dienes method of mathematics learning gives children greater opportunity to know each other during the educational process and within a more unified social structure than the traditional methods.

9. The students learning mathematics with the Dienes method have a more favourable attitude toward the mathematics lesson, and probably the whole of their school activities, than the students learning mathematics in the traditional way.

10. The Dienes method of mathematics learning decreases the students' anxiety.

11. The Dienes method of mathematics learning increases the students' achievement motivation.

12. Learning mathematics with the Dienes method causes a favourable change in the students' personality.

13. The Dienes method helps students to face problems rather than to avoid them.

In the New York evaluation study the fundamental question was: Is there any difference in the examined areas between the children in grades 4 and 5 who belong to a private school in New York City and those who belong to a public school in Sherbrooke, if both have used the Dienes method of mathematics learning?

Five particular questions were of interest:

I. What is the level of knowledge in "traditional mathematics" of the Fleming school children in comparison with those of the Ecole Eymard ?

II. Is there any difference in the achievement in the field of "Modern" Mathematics between the children of the two schools?

III. How well are the children of the Fleming school doing on a test which is designed to measure learning ability in comparison with those of the Ecole Eymard?

IV. What is the attitude of Fleming school students towards their school, Math lesson and their teachers?

V. What is the position of Fleming school students in their attitude towards problems in comparison to students in Sherbrooke (Canada) and in Cedar Falls, Iowa (U.S.A.).

13.4 Statistical design

In this section, we shall present the statistical design of our study by answering the questions raised by Sitgreaves (1964) to "every statistical investigator".

13.4.1 Population of interest

What categories of individuals (or elements) do we want to obtain knowledge about? In statistical terms, what are the populations of interest?

In our evaluation the two populations of interest are those elementary school children who have learned mathematics with the Dienes-method more than two years ago and those who have learned mathematics with a "traditional" method. (For the difficulties in defining "traditional" methods, refer to section 8.6 in Vol. I.)

13.4.2 Selection of groups

How are these individuals (or elements) to be identified? What groups can actually be studied?

In the three studies described in detail, we examined children from experimental and control classes from a Hungarian school, children from an experimental and from a control school in Sherbrooke, Quebec, Canada, and children from a New York City school (see Vol. I, chapters 7, 8, 9 and 10 for the description of these schools, and chapter 14 for a more detailed description of the groups studied.)

13.4.3 Variables of interest

What attributes of the individuals (or elements) do we want to study? That is, what are the variables of interest? Frequently, we are interested not only in attributes at a single point in time, but also in the manner in which they change or develop over time, particularly when some specified treatment or condition is introduced during the time period under study.

In most evaluations, it is possible to differentiate between three different types of variables:

(i) Independent Variables (The treatment of the groups.) In our case, it was a binary variable, experimental-control conditions in the Budapest and Sherbrooke study and French Canadian versus American background in the New York study.

(ii) Background Variables

Background variables are the kind of variables which have more or less static characters, as age, sex, I.Q. level, socio-economic status, etc. The background variables studied in any experiment are selected usually because they seem pertinent to the investigation...

In general, these variables need to be selected thoughtfully and should not be included indiscriminately. For some background variables, such as age and sex, the question of measurement or of categorization is easily solved. For other variables, such as I.Q. and socio-economic status, suitable measuring instruments must be found. (Sitgreaves, 1964).

Background variables can be used mainly for two purposes:

a) Defining interactions (In this case, the background variables are usually called "stratifying" variables.) By this we mean that the experimenter may be interested in seeing whether the learning process under study is different for boys and girls, say, or for children of different I.Q. levels.

In fact, one thing we wanted but were unable to accomplish was to see whether different educational environments have a different effect on each group of students (e.g., on boys and girls in different age groups; children rated as high, medium, and low on certain dependent variables* in different age groups; children rated as high, medium and low in different combinations of dependent variables).

There were numerous instances when successful attempts have been made to show the relation between program effects and greater or lesser ability of the individuals as measured by standard tests of intelligence. It is quite possible that a certain programme A has a greater effect on a group X (e.g., on brighter students) than on another group Y (on less intelligent students) (see Hovland et al., 1949).

This differential effect as a function of stratifying variables, such as I.Q. or other measures of ability, is of the form of interactions talked about as significant effects on the statistical tests. It is also possible (but there is much less evidence in the literature) that a certain method A is superior for Group X while a method B is superior for Group Y. In fact, the existence of this kind of "reversal" interaction (that is, one programme is better for certain persons whereas other programmes are better for other persons) is a main argument in the whole movement for the individualization of instruction. To examine this in the frame of our investigation would have been a very interesting possibility.

---

* Of course in these cases, these dependent variables would be used as "stratifying" variables.

b) Matching purpose

"In much of curriculum experimentation, as in other scientific research, we are interested in making comparisons among groups. This may be less in evidence in the development of a completely new course where the experimental process is evolutionary and the primary purpose is exploratory. Even here it may be desirable to make comparisons of groups in terms of background variables such as sex and I.Q. In many cases of curriculum experimentation, however, the matter of primary interest is the comparison of experimental and control groups. In evaluating the effectiveness of a particular experimental program by comparing the results of experimental and control groups, it is vitally important to be certain that any observed differences are due essentially to the difference between the experimental and control programs, and not to other factors which may affect the learning process. In planning these experiments, every effort should be made to ensure that, except for the particular experimental variables under study, all the groups in the experiment are as comparable as possible with respect to other pertinent factors. This can be achieved partly by the precise definition of the criteria by which the groups of interest are identified. In addition, if there are large numbers of individuals available for study, it may be possible to set up classes which are matched on some of the pertinent background variables. Thus, it may be possible to schedule classes for the different experimental conditions at the same time of day.

It is obviously impossible, however, to match all the pertinent variables. For example, an extremely important factor in any classroom situation is the teacher. While general criteria for teachers in the experiment can and should be set up, experience has shown that teaching effectiveness within a program tends to vary, and that the same teacher may not be equally effective in two different classroom situations. In designing the experiment it is important, therefore, to introduce the two Rs of experimentation: randomness and replication. Randomization should be used wherever possible, so that the effects of the uncontrolled variables tend to balance out in the long run. Thus, if classes within a particular school district are to be set up to participate in an experiment, it may be possible to form the classes by randomly selecting students from the available supply, to assign teachers to the classes in a random way, and to use randomization in determining the particular program to be followed in each class. Replications of the experiment should be carried through to obtain some measures of the variation in the results which are introduced by some of the uncontrolled variables. Thus, replications of the same program in different classes within the same school should enable one to measure the variation in results attributable to the difference between teachers. Replication of the program in different schools within the same area should lead to a measure of interschool variability (Sitgreaves, 1964)."

In our study, unfortunately, we were not able to introduce either of the two "R's". However, we tried to match experimental and control groups as well as we could (see chapter 14).

(iii) Dependent Variables

These are dynamic variables, with which the experimenter hopes to find significant differences.

"The problems of measuring dynamic variables which are associated with the learning process and of evaluating changes in an individual over time are generally much more difficult. In many areas of learning, pencil and paper tests which are useful for measurement purposes have been, or can be developed. In some learning situations, however, for example, in learning to speak a foreign language, ratings of performance are needed. When ratings are given by teachers or others, care should be taken to develop objective criteria on which the ratings can be based. In the discussion, we will use the term "test" to include any suitable evaluative procedure.

Consider now a curriculum experiment in which a new course has been introduced for the first time. Tests which measure the progress of individuals taking the course are needed, and these tests must be suitable not only to the material in the course but also to the individuals being tested. Tests prepared for other learning situations, or for other groups, may not prove particularly useful in evaluating the progress of the experimental group except when it is desired to compare the performance of the experimental group with that of other groups. For example, in teaching logic to academically talented fifth graders, it would seem desirable to develop some tests particularly for fifth graders. On the other hand, if the fifth-grade course material is comparable to material in a course currently being taught to college freshmen, it may also be of interest to compare the performance of the two groups on the same test (Sitgreaves, 1964)."

The dependent variables involved in our studies are shown in Table 13-1. As we described in section 12.2.2, we were intending to study many other variables, but because of technical limitations we had to choose only the most promising ones.

13.4.4 The problem of measurement

How are these attributes to be measured? How are changes over time to be evaluated quantitatively? What are the variables we actually observe? How are they related to the latent variables of interest?

At present, we can report results gained at a single point in time but as we plan to repeat at least some of the measurements, we hope that in the future we shall be able to study the changes in these variables. In fact, although of course with less reliability than from a real longitudinal study, we hoped to gain some information concerning the changes over time with groups of 3, 4 and 5 years of exposure to experimental conditions.

Table 13-1 shows the tests we actually used in our studies. With the statistical analysis of the data, we always supposed that these test really measured what they were intended to measure.

13.4.5. Sampling

How many individuals can be studied? Is this the totality of individuals of interest to the experimenter or does it represent only some fraction of that group, i.e., a sample? If the experiment is carried on with samples, how are the samples to be chosen? More than this, how is the experiment to be designed?

We could measure only a small fraction of the total population. Our experimental school was chosen in the Budapest study because it was judged by the educational experts as a "typical" one, while in the Sherbrooke study, the choice was made because it was the most closely connected to the origin of the teaching method. The experimental groups in both cases were chosen by a matching procedure (see chapter 14).

The Fleming School in the New York study was chosen because of a practical reason: the director of the school felt the necessity of an evaluation of the effect of the Dienes method.

13.4.6 Statistics

Once the experiment is in progress and data are collected, how are these data to be summarized numerically? What kind of descriptive statistics are needed? What is the meaning of the descriptive statistics for the group or groups that have been studied?

Although in some hypotheses, only descriptive statistics were used (averages, graphs, etc.), the meaning of which were merely for demonstration purposes, most of the statistical treatment of data in our studies can be divided into two main parts: a) Testing for significant differences between the experimental and control classes in the different measurements (including subtests); and b) Looking for correlations between the tests (and some of the subtests).

In both cases, it seemed to be preferable to use non-parametric statistical tests, with the exception of these tests for which we may have supposed a normal distribution. Some of the tests required a special statistical treatment before entering this general picture: the factor analysis on the results of the attitude test (see section 21.4.2) and the item analysis for the Sherbrooke Problem Facing Questionnaire (SPFQ; see 23.2.5).

Because of the lack of standardized tests and because of the relatively small samples which did not give enough information about the distribution of the data we used non-parametric methods often. Statistical significance was always determined at the 5% level. Although in some hypotheses we had clear expectations as to the direction of the change, in the interests of fairness we used two-sided statistical tests in all cases.

One of the great problems in designing an appropriate evaluation is that the data we gathered and the quantitative methods we use for analyzing them should enable us to measure the cause-and-effect relations existing among individuals, instructional environments, and learning.

"We need data and procedures to estimate the relationships between the naturalistic environments of learning and the learner's intellectual and social processes, on the causal side, and the learning, on the effect side. That is, in evaluating instruction, we are usually trying to estimate cause-and-effect relations in non-experimental data to make judgements and decisions about the instructions." (Wittrock, 1970; p. 3)

In doing so, one would need a fairly sophisticated statistical design, one which unfortunately could not be afforded in any of the three studies reported here. For example, it seemed necessary (as is usual in this kind of educational study; Bock, 1966) to use multivariate methods in order to handle the complex set of data; a multivariate analysis of variance (Jones, L.V. 1966) to test for differences or a factor analysis (see e.g. Harman, 1967) to find connections among the variables.*

13.4.7 Generalization

If we have been studying samples from the classes of interest, how do we generalize from the data for the samples to the totality of individuals of interest? What techniques of statistical inference do we need? How are these generalizations to be interpreted?

---

* A multivariate analysis of variance could have been carried out on the test scores, if we "believe" that the tests used are measuring what they intend to measure or on the factorial scores (derived from a factor analysis carried out on all the subtests) if we "believe" that these factorial scores would represent more clearly interpreted variables.

Because of the selection of schools, there is not too great a possibility of generalizing our results except perhaps in the case of the Budapest study for other Hungarian schools involved in the experiment.

The field where we think generalization is possible is in the methods we used in our evaluation. As we already mentioned before, our hope is that these evaluations will form a basis for further and more representative cross-national evaluations in the field of learning in modern mathematics.

Table 13-1 Theoretical Model

Objective	Environment	Personality factor	Criteria		
			Budapest Study	Sherbrooke Study	New York Study
To endow children with a reasonable mastery of important techniques and knowledge of Mathematical facts.	A by-product of activities involving different kinds of problems.	Knowledge level in "traditional" maths.	An achievement test based on the "traditional" maths curriculum	Iowa Tests of Basic Skills, Form 5, Test M (Mathematics Concepts and Mathematics Problem Solving; Hieronymus and Lindquist, 1971.	
To create a good understanding of basic mathematical concepts (e.g. variable, function, isomorphism, linearity, etc.) and of their inter-relations.	Problems of set, logic, topology, combinations, probability, statistics, groups, etc. Using the "multiple embodiment" technique, etc.	Mathematical ability.	An achievement test based on the experimental maths curriculum	An achievement test based on items from the Sherbrooke Item Bank (Connes, 1973)	
To prevent fatigue in the children as a result of the maths lesson.	Less external pressure to achieve; fewer restrictions. Avoidance of monotony by different kinds of tasks, materials, activities and interesting problems.	Fatigability.	Reaction to visual and acoustic stimuli; heart rate; visual field; finger tremor.	╳	╳
To capture the child's attention.	Variety of materials and activities. Problems interesting to children.	Attention level.	Toulouse-Pieron test. (Pieron, 1929)	╳	╳
To develop mathematical learning ability.	"Learning centred" rather than "teacher centred" arrangement (independent, individualized learning).	The dynamic of mental development (Vygotsky, 1962) Learning ability; something determining the performance and the speed of learning in a learning task, as well as the creative application of the gained information.	Secret Writing. A test of information acquisition and creative application (IA) Hungarian form (Klein, 1970).	French form (Klein, 1972c)	English form (Klein, 1973f)

Objective	Environment	Personality factor	Criteria		
			Budapest Study	Sherbrooke Study	New York Study
To help the child explore on his own his original "creative" ideas. To allow the child's special "divergent thinking" to blossom.	Letting the children discover regularities. Letting the children construct tasks themselves. Giving tasks with a variety of possible solutions (indefinite tasks, open questions, etc.).	Creativity: fluency (ability to give multiple answers to the same given information); flexibility (spontaneous shift of set), originality (rarity of responses in the population) (Guilford and Hoepfner, 1971).	Items from Torrance's (1962, 1966) creativity tests.	Tests de Pensée Créative. Forme A (Torrance, 1972)	X
To decrease rigidity of thinking.	Sensorimotor experiences with a sufficient number of varied, concrete materials. A great variety of working tools and problem situations.	Perseveration; rigidity: "A tendency to persist in responses that may previously have been suitable in some situation or other but that no longer appear adequate to achieve current goals or to solve current problems" (Wesley, 1953).	Test de persévération (Zazzo and Stambak, 1969)		X
To help the children know each other's abilities better.	Group activities (e.g., competitions between groups, grouping the children to find the solutions of a more difficult problem). Right to "cheat" when problems appear. Giving more emphasis to intrinsic motivation (not giving reasons for envy, arrogance, contempt, etc. which could distort the children's opinions about each other.	Mutual appraisal: the capability of evaluating other members of a group.	"Who are the 3 most clever pupils in your class?" - a tailor made test	X	X

Objective	Environment	Personality factor	Criteria		
			Budapest Study	Sherbrooke Study	New York Study
To create a feeling that mathematics is "fun" or "exciting" or "challenging" or "rewarding" or "worthwhile" (Davis, 1964).	Intrinsic rather than extrinsic motivation. Playing at every level. Giving the children the pleasure of discovery as often as possible. Providing a variety of interesting working (playing) tools and problems. The practice of the "deep end" principle.	Attitude: "the evaluative dimension of the total semantic space" (Osgood et al., 1957).	Semantic differential and attitude measurement (after Osgood et al., 1957).		
To decrease the anxiety aroused in the child by the mathematics lesson.	Intrinsic rather than extrinsic motivation. (No emphasis on grades, no punishment, but interesting problems, etc.) Never pushing the children to work faster. Right to make mistakes. Giving the children many possibilities to prove to themselves that they are able to solve problems alone. Right to be different from others.	Anxiety.	Children's Manifest Anxiety Scale (CMAS; Taylor, 1953; Hungarian Version: Forrai, 1968).	Epreuves d'Anxiété (Demangeon, 1966)	✗
To develop in the child a strong motivation to achieve according to his abilities.	Group activities (e.g., competitions between groups). Independent learning. Interesting problems according to the student's need. More emphasis on intrinsic motivation (e.g., joy of discovery) than on extrinsic motivation (e.g., marks).	Achievement motivation (need to achieve): "to overcome obstacles, to exercise power, to strive to do something difficult as well and as quickly as possible" (Murray, 1938).	Thematic Apperception Test (TAT; Murray, 1943).	✗	✗

Objective	Environment	Personality factor	Criteria		
			Budapest Study	Sherbrooke Study	New York Study
To develop a positive attitude toward problems.	Indefinite problems. New problems each time.	Attitude	X	Sherbrooke Problem Facing Questionnaire. French form (Klein, 1973 h)	English form (Klein, 1973 h)
Help to develop the whole of the students personality.	Freedom in activities: abundance in the possible activities	Reserved - Warmhearted. Less Intelligent- More Intelligent. Affected by Feelings - Emotionally stable. Phlegmatic - Excitable Obedient - Assertive Sober - Happy-go-lucky Expedient - Conscientious Shy - Venturesome Though-minded - Tender-minded Vigorous - Circumspect Forthright - Shrewd Self-assured - Apprehensive Casual - Controlled Relaxed - Tense Anxiety Extraversion	X	Children's Personality Questionnaire Form A and B (Porter-Cattell 1963).	X

Chapter 14

The structure of the three evaluation studies

14.1 The Budapest study

14.1.1 Purpose of the study

This study was designed to fulfill three major objectives:

TO ANSWER SOME QUESTIONS RAISED BY THE ADMINISTRATIVE EDUCATIONAL LEADERS AND PARENTS.

Under this title we had to answer three questions:

a) Are the pupils in the pilot classes at least on the same level in the traditional subject matter (which they learn during a small fraction of their time) as the pupils who learn the same material for the entire school year?

b) Did the pupils in the pilot classes gain any knowledge of such "new" topics as sets, logic, functions, etc?

c) What is the price of any advantage in learning, in terms of such factors as pupil fatigue?

TO ANSWER SOME QUESTIONS RAISED BY THE TEACHERS, EDUCATORS, AND PSYCHOLOGISTS INVOLVED IN THE EXPERIMENT.

Everyone involved in this experiment was delighted with the "personal aims of mathematics teaching" as stated by the project: "towards the use of mathematics learning to contribute to personal fulfill", or as Dienes expresses it in his Building Up Mathematics (1960), "building up personality". We can agree with him that "it is clearly the educator's business to ensure that children get every chance of achieving as full a degree of integration as they are capable of, just as it is his business to ensure that children reach their fullest potentialities in the intellectual sphere." (Dienes, 1960; p. 25) The question the teachers, educators, and psychologists wanted answered was whether the OPI method had any success in developing the students' personalities in both cognitive and orectic regions (Dienes, 1959a); did it affect, for example, the students' learning ability, creativity, flexibility, or their attitude towards the mathematics lesson, their anxiety or their achievement motivation?

To make this distinction between the two sets of questions does not mean that the teachers, educators, and psychologists were not interested in the students' knowledge, not that the administrative educational leaders were not interested in the students' personality development. It means only that to find the answer to the first type of questions is mainly an external necessity, while to find the answer to the second type of question is mainly an internal necessity.

TO START BUILDING AN INSTRUMENT SENSITIVE ENOUGH TO MEASURE THE PSYCHOLOGICAL EFFECT OF THIS NEW KIND OF LEARNING SITUATION.

We hoped that while we were trying to find the answers to the first two (and especially the second) groups of questions we would collect testing instruments which could serve as a base for future evaluation processes. In this sense we wanted to treat our data not only as the examination of the experimental effect, but also as the examination of a new evaluation battery.*

---

* For more information on the Budapest study and its results, we may refer the interested reader to Klein (1972a, 1973c, 1973d).

14.1.2 Organization

SCHOOL

For this study we chose a school in Budapest in an environment of below average socio-economic level.

This school joined the scheme in 1965 with a class of first graders (6^+ years of age). This class was in grade 5 at the time of the evaluation and four other pilot classes were in grades 4, 3, 2 and 1, respectively, each having started with the experimental method in grade 1.

CLASSES

We used the grade 2, 3, 4 and 5 pilot classes as experimental groups and four parallel classes from the same school as control groups. All the tests were done by every student, except the measures of fatigue, which, because of the limited time, were performed by only twelve randomly selected students from each class. The number of students in each class according to their sex is shown in Table 14-1.

In practice, because of illness and some other reasons we lost a few students in most of the tests, which caused slight differences in the number of persons taking part in the experiment test by test. Because some of the tests were inappropriate for younger children, they were carried out only in grades 3, 4 and 5 and one (the measurement of learning ability) only in grades 4 and 5.

Table 14-1

The number of students in each class according to sex in the Budapest Study.

Treatment / Grade	Experimental Classes			Control Classes		
	Boys	Girls	Σ	Boys	Girls	Σ
2	9	20	29	15	14	29
3	18	9	27	10	18	28
4	14	10	24	14	15	29
5	18	13	31	12	12	24

Table 14-2

The average IQ's in the experimental and control classes of the Budapest Study.

Classes	3rd		4th		5th	
	experimental	control	experimental	control	experimental	control
I.Q.	96.1	94.5	107.8	106.7	107.4	105.4

The teachers both in the pilot classes and in the control classes were "common soldiers" of education. They teach every subject (except in grade 5), not only mathematics. None of them is specially qualified to teach mathematics. They were rated by their superior as "average teachers". (In fact, to be a "good teacher" has different meanings in the traditional and in the experimental teaching conditions: e.g., the autocratic teacher who is rated superior in traditional conditions could fail in the experimental ones, and the "democratic" teacher who is rated poor in traditional conditions could be very successful in the experimental ones.)

TESTING PERIOD

The measurements of fatigue were taken at 8 a.m. (before the mathematics lesson) and at 9 a.m. (after the mathematics lesson). Every other test was given only once, between 8 a.m. and noon. The entire experiment was carried out from September to December 1969.

MATCHING

The students in our study were not selected in any way, so we have supposed that the experimental and control classes were equal on each of the parameters at the beginning of their studies. Because all the classes were chosen from the same school milieu, we can suppose that all differences found at the time of the examination are due to the only respect in which their school environment differed, viz., mathematics teaching.

We examined this assumption on only one parameter, by comparing the results of the experimental and control groups on a "traditional" intelligence test: the Hungarian version of the Stanford-Binet test. The average I.Q.'s are shown in Table 14-2.

None of the differences between the experimental and control classes was significant at the 5% level using the Mann-Whitney test.

In recent years there have been some interesting experiments to explore the relationship between the learning of modern mathematics and performance on intelligence tests.* If, as some of these experiments suggest, intelligence test performance is boosted by experience of modern mathematics, then parity in performance after experimental and control treatments would imply that the intelligence level of the experimental group had originally been lower. In this case, results which did not favour the experimental group would be equivocal.

14.2 The Sherbrooke study

14.2.1 Purpose of the study

The objective of the Sherbrooke study was to build a set of instruments sensitive enough to demonstrate the various expected effects of the extensively-used Dienes method of mathematics teaching, and, more generally, to build a set of instruments capable of measuring some of the human factors which are so much affected by a pedagogical system. The two practical tasks necessary to the realization of this objective are immediate derivatives of these two problems:

---

* Investigations of this relationship have yielded contradictory results: Hull (1963) and Fogelman (1968) reported significant differences between the experimental and control classes using the Lorge-Thorndike test. Williams (1972b) found no significant differences using the same test (except in one subtest), nor in the N.F.E.R. Picture Test (Williams and Fogelman, 1972). Fogelman (1968) and Williams (1972b) did not find significant differences using the Raven Matrices (except the latter in one subtest). We hope to carry out further investigation of this topic in the near future, but, to date, results neither support nor contradict our use of the traditional intelligence tests as a matching variable.

1. the selection of standard tests, among those which are available, appropriate to the expected outcome of the Dienes method; and
2. the adaptation of some standard tests, and the development of some new tests, designed to measure important outcomes in the fields where standard tests are not available.

The method chosen to carry out the two practical tasks consisted of a kind of combination of evaluation and test development effort. First, traditional standard tests were employed in the evaluation process to find differences between experimental and control classes. Second, these same tests and the differences found with them were used to develop the construct validity of the new tests. Thirdly, these new tests were incorporated into the evaluation battery. To break out of a seemingly vicious circle, one of the new tests was validated on grade 5 children at Cedar Falls, Iowa, U.S.A. (see 23.2.5) as well as in the Sherbrooke schools. The design of this study is quite similar to the Budapest study (see 14.1), mainly to facilitate the comparison of these results with those obtained in the Hungarian study. Recently, cross-national education has become a topic of general interest (see e.g. Foshay, 1962; Husen, 1967; 1969; Miel and Berman, 1970). Because the Dienes method is already used widely throughout the world, it seemed natural to further develop it as an instrument which can be used in several countries.

14.2.2 Organization

The testing was carried out between May 11 and June 21, 1972, except for the achievement tests, which were performed in April and May, 1973. It included one class each from grades 3, 4 and 5 (whose students were completing their third, fourth, and fifth years with the Dienes method, respectively).

Finding the control classes was not as easy as in the Hungarian study: there were no classes unaffected by the experimental method in the same school. In order to choose control classes from another school, experimental and control classes had to be matched in all other aspects than the one we wanted to study: the learning of mathematics. The methods for matching the control classes were the following: First, the Commission des Ecoles Catholiques de Sherbrooke was asked to name the schools which they felt used similar teaching techniques and facilities, and in which the students had similar socio-economic backgrounds to those in the experimental schools. From the possibilities provided by the commission, the school we thought best fitting these conditions was selected. Our choice was verified in only four factors:

i. Social composition The children were divided into four groups according to the social status of their parents, and χ^2-tests were computed to see whether the distribution of the students in these four categories was the same in the experimental and control classes in each grade separately (see Table 14-3).

ii. Intelligence level We used Lorge-Thorndike intelligence test scores obtained when the students began grade 1, as a measurement of the students' original intelligence status. t-tests were carried out to see whether the average intelligence of experimental and control children was different or not at the time when they began their schooling (see Table 14-3).[*]

Three different intelligence tests were carried out during our evaluation: the Raven (1958), the Otis-Lennon (1967a, b), and the Cattell Culture Fair (English, 1949; French, 1967). Separate and combined scores from these tests were to be used as part of the matching procedure.[**]

iii. Conceptual level The students' conceptual level was supposed to be measured by a test developed by Hunt (1970b)and adapted to the Quebec population by us). This measure was planned to not only match experimental and control classes in this factor, but also to use as a "stratifying variable" (see 13.4), that is to arrange the students into low, average, and high conceptual level groups in order to examine the effect of the Dienes method in the three groups.

Because of technical difficulties the students' answers on this test were not scored.

iv. Teachers' attitude An attempt was made at matching teachers' attitudes, using the Minnesota Teacher Attitude Inventory (Inventaire Minnesota des Opinions de l'Enseignant; Cook et al., English, 1951; French 1972). In the present evaluation, matching in this factor was not entirely successful (see Table 14-4).

---

[*] Tables 14-3 to 14-5 will be found at the end of this chapter.

[**] In fact the main objective to include these three intelligence tests into our evaluation model was to reexamine the question whether or not intelligence test performance can be boosted by modern mathematics teaching. Ohlsen (1963, p. 270) writes that "since 1920, we have learned that mental ability measured by tests is not stable under all conditions" and then - based on a paper from Findley (1961) - he states quite strongly that "a learning climate that challenges and stimulates, but does not frustrate a child tends to increase I.Q., while a cold, forbidding educational climate operates in reverse for many children." But in fact the research data are quite contradictory in this field. (Williams and Fogelman, 1972; Fogelman, 1968; Williams, 1972b, Jensen, 1969b.)

In order to match intelligence and socio-economic status, some children were omitted for considerations of experimental-control differences. After this correction there were no significant differences in these two factors (see Table 14-3).

14.3 The New York Study

14.3.1 Purpose of the study

The New York study consisted of a project with the objective of making an evaluation between two schools using the Dienes method of teaching modern mathematics at the primary level but in two different cultural and language backgrounds: the experimental school in Sherbrooke described in Chapter 9 Vol. I and a private school in New York City described in Chapter 10 Vol. I.

The original plan had been to evaluate the various expected psychological "side effects" of the Dienes method of teaching modern mathematics according to the pattern followed in the Sherbrooke and the Budapest Study, i.e. testing and comparing children from an experimental school (with the Dienes method) and from a control school (as similar as possible to the experimental school, but with a traditional method of teaching mathematics).

Unfortunately the permission to examine children from the control school was withdrawn at the last minute. Because of this, the evaluation changed its direction: instead of looking for differences caused by the different educational environment, we were now looking for differences caused mainly by the different socio-economical, cultural and language backgrounds. With both schools using the Dienes method, we also intended to find out how well this method would work in the two schools.

14.3.2 Organization

Only aspects concerning the New York School are dealt with here, because comparison with the Sherbrooke School is based on data collected within the Sherbrooke Study. For organizational aspects concerning that school, see under 14.2.2.

Testing Period - The testing was carried out from April 10th to April 20th, 1972. Those tests which needed a higher concentration (knowledge in modern and traditional maths.) were passed in the earlier periods of the morning or afternoon sessions.

Schools and classes The school in New York is described in detail in Volume I.

The children examined belonged to two classes at the grade 4 level and two classes at the grade 5 level which had been affected by the Dienes method for two and three years respectively. Those children who entered the school after 1970 were left out. Table 14-5 shows the number of students and the average intelligence level (also separately for boys and girls), in each class, each grade and together.

Table 14-3

Basic statistical data of experimental and control groups in the Sherbrooke study.

Grade	Condition	Group	Number of Students	Sex			Social class					IQ		
				Boys	Girls	χ^2 $p = 3.84$	1	2	3	4	χ^2 $p = 7.84$	Average	Standard Deviation	t $p = 2.02$
3	exp	total	25	12	13		10	4	7	3		——	——	
		permanent*	20	10	10	0.01	7	4	7	2	0.69	112.0	14.48	
	contr	total	33	17	16	0.02	12	5	13	3	0.56	——	——	0.237
		permanent*	21	11	10		7	4	9	1		113.14	16.24	
4	exp	total	24	11	13		5	7	5	7		——	——	
		permanent*	20	11	9	0.86	4	6	5	5	1.37	100.05	8.28	
	contr	total	29	17	12	1.21	6	7	10	6	0.10	——	——	1.825
		permanent*	18	13	5		4	5	4	5		106.33	12.69	
5	exp	total	27	12	15		6	4	12	5		——	——	
		permanent**	23	11	12	0.02	6	3	11	3	4.49	106.30	9.45	
	contr	total	28	13	15	0.09	12	6	6	4	2.74	——	——	0.493
		permanent**	21	11	10		8	4	5	4		108.09	14.37	

Notes

* At least 3 years in this school

** At least 4 years in this school

*** (see next page)

the critical p values refer to the 5 percent level of significance

Table 14-3 cont.

***

Description of the categorization of the socio-economical status of the parents (father or mother)

1. Professional status: doctor, lawyer, etc.
   years of education: 14 and more
2. "White collar": office employee, clerks, managers, etc.
   years of education: 11 to 13
3. Skilled worker: plumber, truck driver, etc.
   years of education: 8 to 12
4. Unskilled labourer and those without work (chronically unemployed),
   years of education 0 to 7

Table 14-4

The results of the Minnesota Teacher Attitude Inventory in the Sherbrooke Study

Groups / Grades statistics	Experimental				Control			
	Score			Percentile	Score			Percentile
	"right"	"wrong"	Attitude		"right"	"wrong"	Attitude	
3	90	48	42	30	61	65	-4	8
4	96	43	53	42	83	56	27	24
5	91	41	50	40	86	46	40	28

Table 14-5

The number of students and the average IQ's[1] in the examined classes in the Fleming School, New York.

Statistic	Grade	4									5									Σ		
	Class	Yellow			Red			Σ			Yellow			Orange			Σ					
	Sex	Boy	Girl	Σ	Boy	Girl	Σ	Boy	Girl	Σ	Boy	Girl	Σ	Boy	Girl	Σ	Boy	Girl	Σ	Boy	Girl	Σ
Number of students		10	12	22	9	12	21	19	24	43	7	10	17	7	11	18	14	21	25	33	45	78
Average IQ[2]		119.5	119.7	119.6	119.3	115.5	117.1	119.4	117.6	118.3	125.0	130.9	128.7	123.8	119.0	121.0	124.4	128.9	124.8	121.5	121.0	121.2

---

[1] IQ values are those of school records (Otis-Lennon test taken in December, 1972).

[2] For calculating IQ averages data from 74 students were available.

Chapter 15

Traditional and Modern Mathematics

In this and the remaining chapters of this Volume we proceed to examine in detail the Hypotheses and Questions listed in 13.3. As already stated, these Hypotheses concern themselves with the affective "side effects" of mathematics learning no less than with its cognitive aspects.

Our immediate concern is with the examination of the first two Hypotheses, which compare the experimental and control classes on Traditional and Modern Mathematics as far as the gain of knowledge is concerned.

15.1 Knowledge of Traditional Mathematics

Hypothesis 1: Although the children in the experimental classes spend much less time on the "traditional topics", their knowledge, even in this field, is not worse than their schoolmates' knowledge.

Our discussion will confine itself to the theoretical problems of achievement measurement, on which a very large amount of literature exists.*

15.1.1 Theoretical Background

Achievement measurement in education is probably as old as the idea of schools. Teachers have always been interested in discovering whether their students really "know" what they have been "taught"; parents are interested in learning whether their children are "better" or "worse" in mathematics than the boy living next door; society's educational institutions are interested in selecting the "better" students; employers are always seeking qualified applicants for specific positions, etc.

---

* The two main approaches are the educational behaviourism exemplified by B.F. Skinner's work in the area of programmed instruction (Skinner, 1953; 1957) and the other is the positivistic approach to mental test theory exemplified by Lee J. Cronbach's work on "generalizability" (Cronbach et al., 1963; Rajaratnam et al., 1965; Hively II, et al., 1968).

There are many different methods of achievement measurement from the less sophisticated subjective judgement or self-rating to the most sophisticated item bank method. As in any other kind of testing, the main problem with these methods is their *reliability* and their *validity*.

For example, the unreliability of *teachers' marking systems* (even in so "objective" a study as plane geometry) was demonstrated as early as 1913 (D. Starch and E.C. Elliott, 1913). School marks are clearly dependent on such irrelevant variables as the personal preference for or acceptance of the pupils by the teacher, or by the sex of teachers or pupils, etc. (Hadley, 1954).

Self-rating also proved to be a function of many factors, e.g., sex (Stier, 1948); socio-economic status (Grossman and Wrighter, 1946; Neugarten, 1946); level of aspiration (Mitchell, 1945; and Frank, 1941); etc.

Standardized tests usually have fairly high reliability but mostly their validity is questionable. "In developing standardized tests, the most widely used array of common objectives in schools, where tests are expected to be sold, is selected. In most instances, these standardized tests measure very well the pattern of objectives. However, this pattern does not correspond to the content of the curricula of many classes where the test is administered (The Mathematics Evaluation Materials Package Project, 1972).

The item-bank method (e.g., Connes, 1973) is more or less able to solve these problems, but until it gains wider use, it is much too expensive and complicated.

Teacher-constructed tests are a kind of compromise between the "subjective" and "objective" methods. They typically reflect the emphasis on local programmes, and hence, avoid the validity problem standardized tests often raise, but their reliability is generally much lower. However, teacher-constructed examinations sometimes can be improved through a careful analysis of programme objectives and some trials of test items that correspond to objectives as well as through applying some statistical methods (Lord and Novick, 1968).

In our evaluation studies, we tried out most of the above mentioned methods and we had to face most of the troubles with them as well. Some of the teacher-constructed tests in the Budapest study proved to be too difficult, the standardized achievement test for measuring "traditional mathematics achievement" proved to be inadequate in the Sherbrooke study, and the item-bank did not work in the New York study. In addition, it must be said that we have gained some useful information from these tests in all of the three studies.

Finally, we wish to mention a fairly new tendency in achievement evaluation: The Cross-national Comparison. This trend has been evident especially since 1959, when the I.E.A. (The International Project for the Evaluation of Educational Achievement) began its first pilot project on mathematics achievement (Husén et. al., 1967), and more and more international studies have been completed (Foshay, 1962; Pidgeon, 1958; Kramer, 1959; Husén, 1969, etc.). The achievement measurement in our study is also a small step in this direction.

15.1.2 The Budapest Study

METHOD

To examine the first hypothesis without appropriate standardized mathematics achievement tests, we had to construct tests for each grade Appendix I). Before using them, the tests were tried out in a school not involved in these evaluation processes.

To construct the test items we followed the requirements of the Reformed Plan of Tuition generally operating throughout Hungary. Each of the tests took into consideration the following requirements:

1. number concept, reading and writing the number, place value;
2. the four arithmetical operations;
3. using measuring instruments, changes in unit of measurement;
4. ability to solve some complex tasks.

The fifth task was always a so-called "task with text". This assessment was carried out in grades 2 - 5.

TREATMENT OF DATA

To compare experimental and control classes, the chi-square test was used on the results of each of the five tasks as well as on the total.

EXPECTATION

Hypothesis 1 would be supported by finding no significant differences between the total scores of experimental and control classes. Significantly higher results in the experimental classes would not contradict the hypothesis.

RESULTS

The results of the traditional mathematics examination are shown in Fig. 15-1 through 15-4. It can be seen that the total scores are higher in the experimental classes in grades 2, 3 and 4 but lower in grade 5. Using the χ^2 -test the differences between total scores in the experimental and control classes were significant only in grade 2 ($\chi^2 = 5.262$), in favour of the experimental class.

Considering all four grades examined, the average score was higher in ten tasks in the experimental classes and in the other ten tasks in the control classes.

As can be seen from Fig. 15-4, the traditional mathematics test (with the exception of the second task) was a bit too difficult for grade 5 in both the experimental and control classes, a fact which has perhaps made our results for that grade less reliable.

As a whole, the results seem to support hypothesis 1 in the Budapest study. Although, the children in the experimental classes spend much less time on the "traditional topics", their knowledge even in this field does not seem to be worse than that of their schoolmates.

Fig.15-1 Results of the traditional mathematics tests in the Budapest study
Grade 2

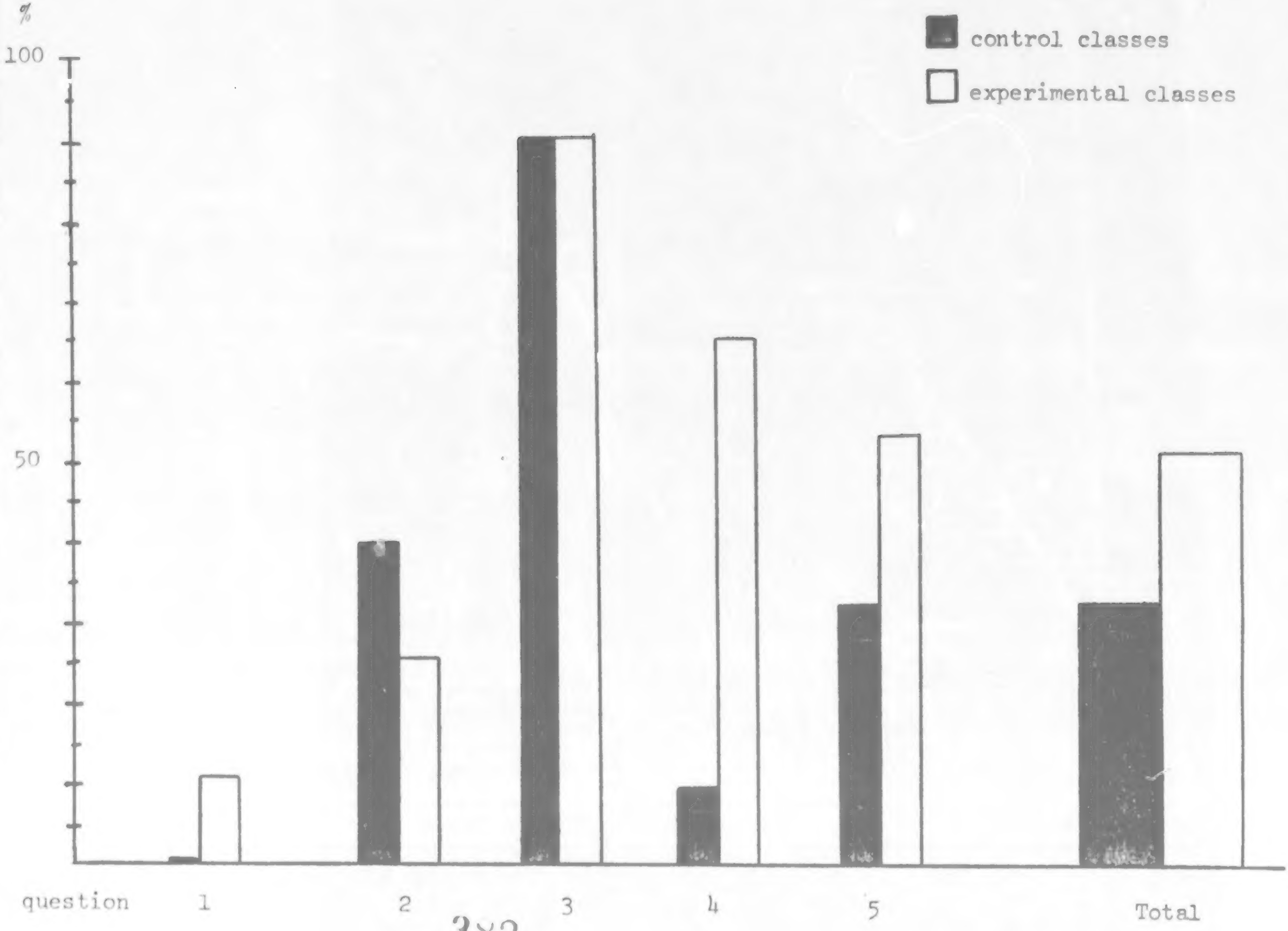


Fig. 15-2 Results of the traditional mathematics tests in the Budapest study

Grade 3

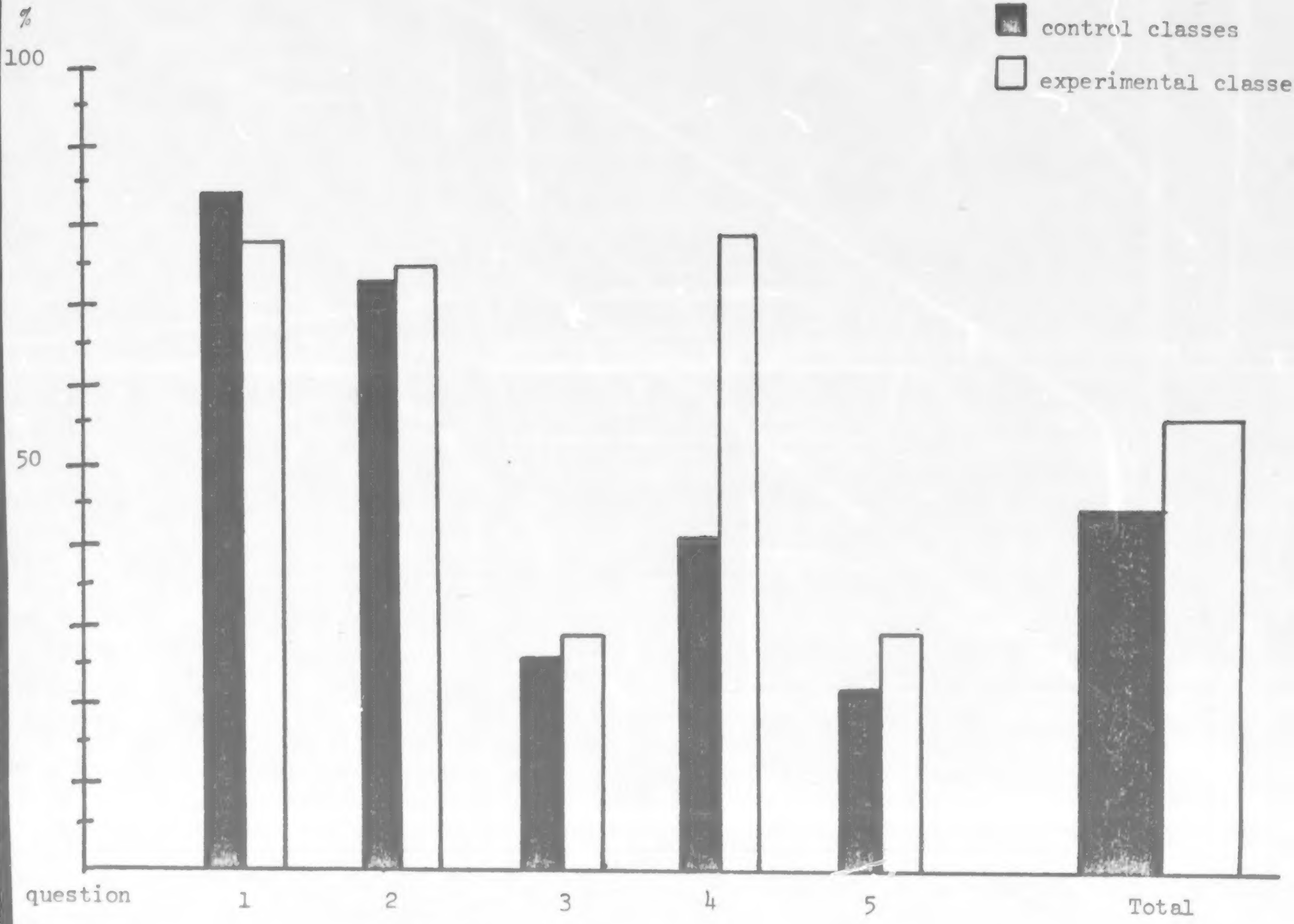


Fig. 15-3 Results of the traditional mathematics tests in the Budapest study

Grade 4

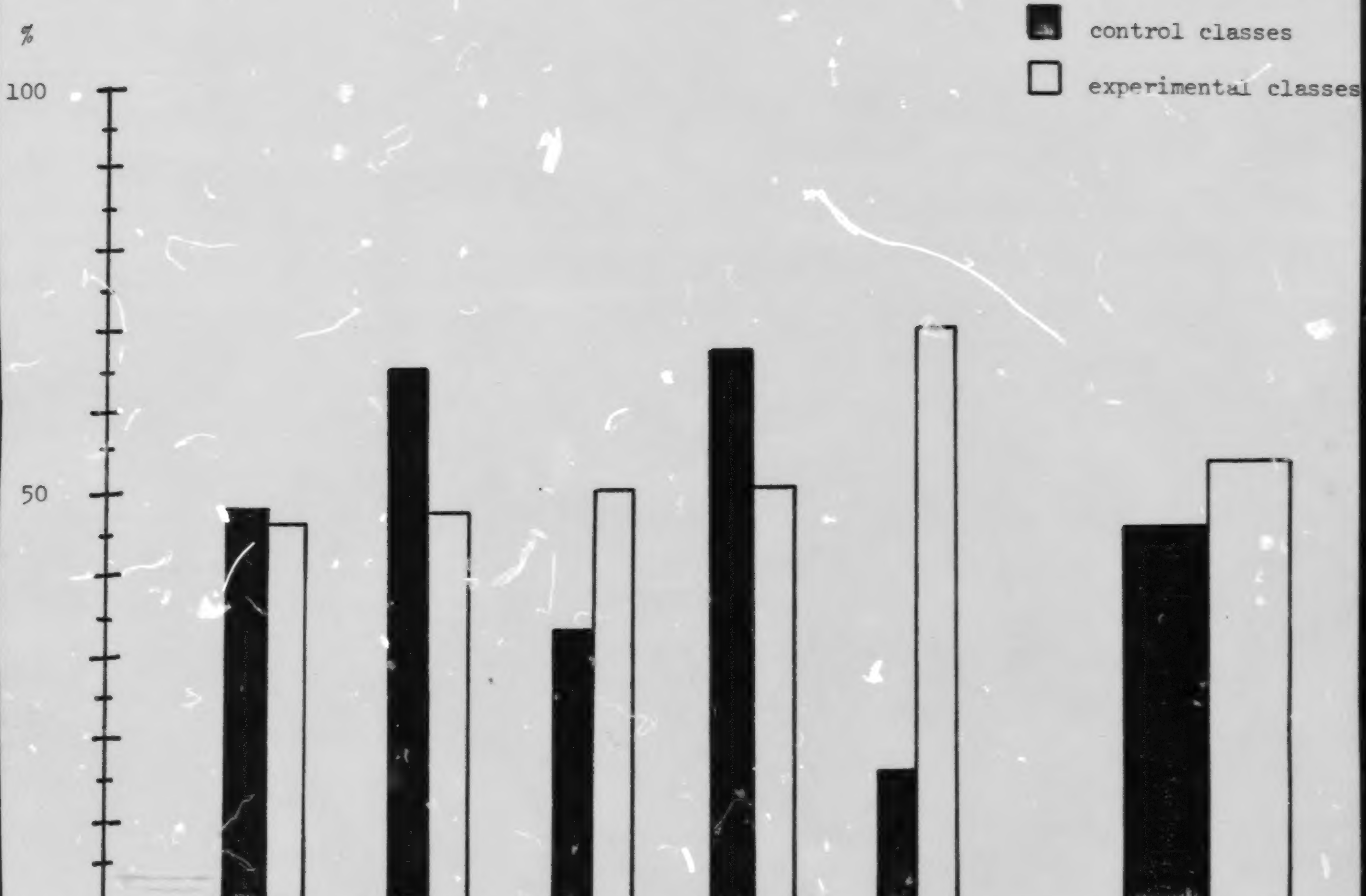
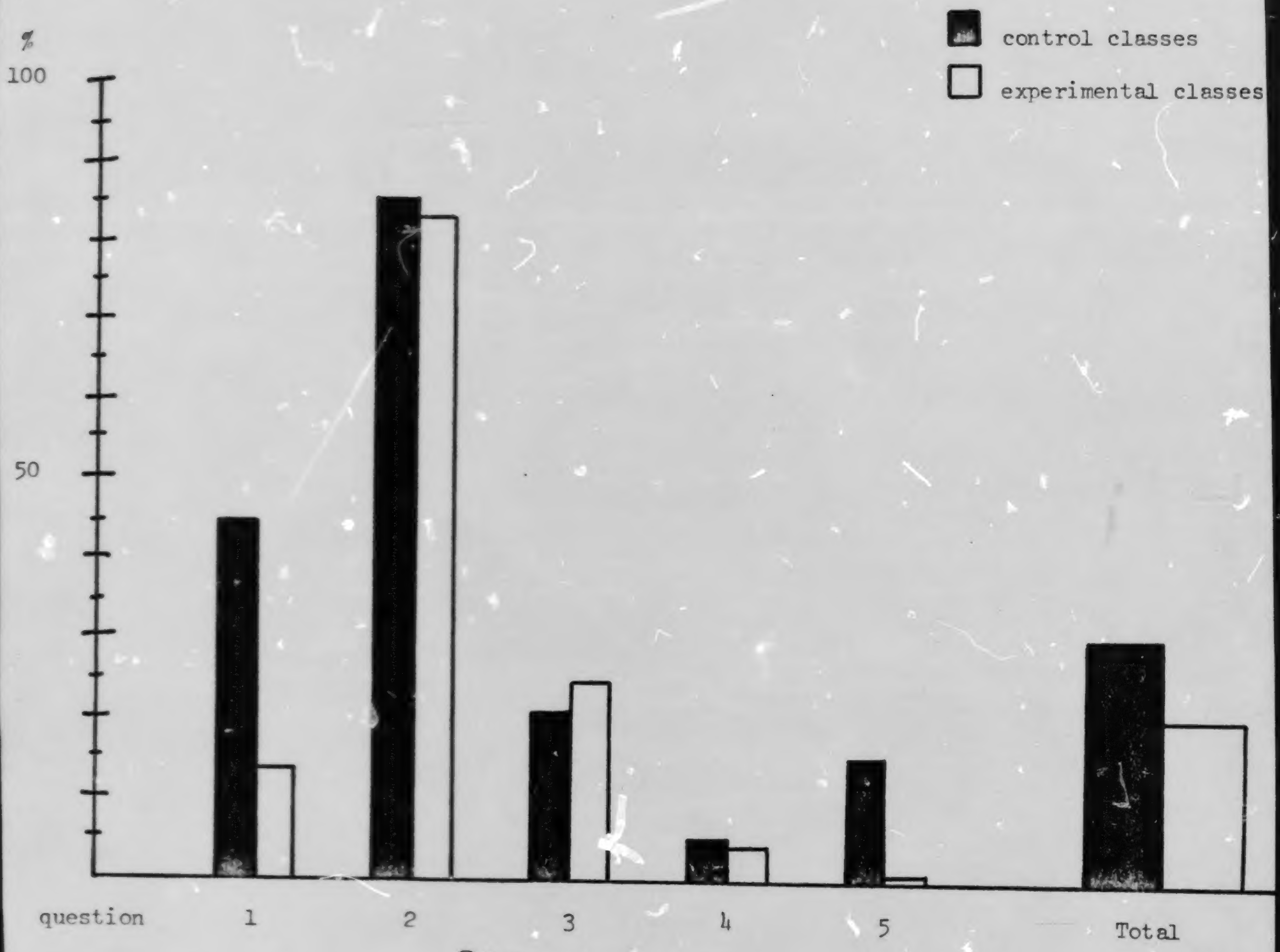


Fig. 15-4 Results of the traditional mathematics tests in the Budapest study

Grade 5



15.1.3 The Sherbrooke Study

METHOD

Instrument: In order to measure children's knowledge in traditional mathematics the achievement test battery "Iowa Tests of Basic Skills, Form 5" (Hieronymus and Lindquist, 1971) was used. For skill in arithmetic two tests are included: Test M-1 "Mathematics Concepts" and M-2 "Mathematics Problem Solving". There is a time limit of 30 minutes given for each of them, which turned out to be insufficient for many students to finish the test.

The tasks of which the tests are made up can be categorized as shown in Tables 15-1 and 15-2 together with the amount of questions referring to them in each grade respectively.

Testing period: The test was done in May, 1973, when the children of grades 3, 4, and 5 were in grades 4, 5, and 6 respectively. The two parts were administered on different days, not in one session as suggested in the test manual.

TREATMENT OF DATA

Comparison between experimental and control groups was made by using t-tests on the number of correctly answered problems in the two parts of the test separately and together.

EXPECTATION

The same as under 15.1.2.

Table 15-1

The number of tasks on the different topics of mathematics in the Mathematics Concept (M-1) subtest of the Iowa Tests of Basic Skills (Hieronymus and Lindquist, 1971).

Categories	Grade 4	Grade 5	Grade 6
Currency	1	-	-
Decimals	-	-	5
Equations	3	4	6
Fractions	4	9	9
Geometry	3	5	6
Measurement	4	4	3
Numeration and Number Systems	8	10	9
Ratio and Proportion	1	1	1
Sets	1	1	1
Whole Numbers	11	8	5
SUM	36	42	45

Table 15-2

The number of tasks on the different topics of mathematics in the Mathematics Problem Solving (M-2) subtest of the Iowa Tests of Basic Skills (Hieronymus and Lindquist, 1971)

Categories	Grade 4	Grade 5	Grade 6
Currency	9	9	8
Decimals	–	–	1
Fractions	1	4	9
Measurement	1	1	–
Per cent	–	–	1
Ratio and Proportion	–	1	2
Whole numbers	16	14	10
SUM	27	29	31

RESULTS

Tables 15-3 and 15-4 summarize the differences in knowledge found in traditional mathematics as assessed using the Iowa tests of Basic Skills: While for the Part "Mathematics Concepts", which stresses more the computational side of mathematic skills, differences are clearly in favour of the control classes; for the Part "Mathematics Problem Solving" (which stresses more understanding and rational use of math processes) the same is true only for grade 5 (equal to grade 6, when the test was taken). However, tendencies in the same direction are found in the other grades as well as will be seen by the t value (significant at the .01 percent level) based on the combined score.

In interpreting the results we must not neglect the following facts: The administered test was given in a translated form which caused some difficulties in understanding for the children; secondly the selection of test items is based on the American curriculum plans and American school books, not on Canadian; thirdly the given time limit of 30 minutes did not allow many children to display their full capacities. Taking this into account, we may conclude that under the above-mentioned test conditions the classes learning mathematics with traditional methods are clearly better in skills concerning computation and mathematical concepts than our experimental classes and somewhat better in mathematics problem solving. We have to leave undecided, however, the outcome if knowledge in this field were assessed under different test conditions, especially without time limits.

15.1.4 New York Study

The question of interest was: what is the level of knowledge in "traditional mathematics" of the Fleming school children in comparison to those

Table 15-3

The mean-differences in the Sherbrooke Study between experimental and control groups in the Iowa Tests of Basic Skills (Form 5, Hieronimus and Lindquist, 1971) Part M-1 Mathematics Concepts.

Groups / Statistics / Grades	EXP			CONTROL			EXP + CONTL				
	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	t	F
5	22.95	6.47	19	28.58	6.89	19	25.76	7.25	38	2.60*	1.14
4	15.92	4.63	13	24.06	6.09	16	20.41	6.82	29	4.09*	1.73
3	14.57	5.85	14	19.79	6.22	19	17.58	6.59	33	2.46*	1.13
Σ^1	-3.35	5.86	46	2.85	6.45	54	0.00	6.91	100	5.03***	1.21

[1] The method of combining score (the meaning of the sign Σ) was the following:

Σ means: $\sum_{i=3}^{5}\sum_{j=1}^{n_i}(x_{ij}-\bar{x})$, where

i = grades
j = subjects
n_i = number of subjects in grade i (only in experimental or only in control group or in both according to the column of the table)
x_{ij} = the score of subject j in grade i
$\bar{x}_i = \sum_{j=1}^{n_i} x_{ij}/n_i$, the arithmetic mean in grade i (only in experimental or only in control group or in both - according to the column of the table)

* significant at .05 percent level

** significant at .01 percent level

*** significant at .001 percent level

Table 15-4

The mean-differences in the Sherbrooke Study between experimental and control classes in the Iowa Test of Basic Skills (Form 5, Hieronimus and Lindquist 1971). Part M-1 Mathematics Problem Solving.

Groups	EXP			CONTROL			EXP + CONTL				
Statistics / Grades	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	t	F
5	8.79	3.35	19	13.11	4.10	19	10.95	4.32	38	3.55**	1.50
4	8.69	3.36	13	8.38	2.55	16	8.52	2.94	29	0.28	1.74
3	13.07	6.15	14	16.63	5.61	19	15.12	6.10	33	1.71	1.20
Σ^1	-1.47	4.52	46	1.25	4.45	54	0.00	4.68	100	3.02**	1.03

[1] The method of combining scores is described in the footnote of Table 15-3

** significant at .01 percent level.

of Ecole Eymard at Sherbrooke?

METHOD

The method of the Sherbrooke Study was repeated. IOWA tests of Basic Skills (Form 5) part M-1 Mathematics Concepts and M-2 Mathematics Problem Solving were used (see 15.1.3).

TREATMENT OF DATA

For each part of the test an overall score of right answers was computed for every student. Comparison between the Fleming school students and the Eymard school students (non permanent students included) were made on these scores, differences in the group means being assessed by the t-test.

RESULTS

In both grades and both parts of the test, the Fleming school students were found to be superior in their knowledge of "traditional" mathematics. Tables* 15-5 to 15-7 show that the differences are significant at least at the .01 percent level of confidence.

The differences between the parallel classes are not consistent in both parts; while for "Mathematics Concepts" the "red" classes are somewhat superior to the "yellow" classes (especially in grade 4) for "Mathematics Problem Solving" the reverse is true.

15.2 Knowledge of modern mathematics

Hypothesis 2: The children in the pilot classes gain a great deal of knowledge in the more and more important field of modern mathematics.

---

* Tables 15 - 5 to 15 - 11 will be found at the end of this chapter.

15.2.1 Budapest Study

METHOD

To examine the second hypothesis we constructed a special achievement test.

TREATMENT OF DATA

To examine the second hypothesis in the Budapest study (as in the other two studies) we used only descriptive statistical evidence. As it would be silly to give a test of modern mathematics to children who had never learned it, the tests of the second hypothesis were carried out only in experimental classes. Tests were performed in grades 2, 3 and 5. Although a test was prepared for grade 4, we were not able to administer it before the end of the school term.

EXPECTATION

Reasonably high scores in the experimental classes would support the hypothesis. Because of the wide variation of learning speed tolerated by the project, and because the test questions were new and therefore required transfer from quite distant learning, we could expect a large dispersion in the test scores.

RESULTS

It seems difficult to demonstrate the hypothesis that mathematical knowledge is increased, but most of our results seem to support such a hypothesis, as they give evidence of a certain level of modern mathematics knowledge.

The results of tests of hypothesis 2 in grades 2 and 3 are shown in Fig. 15-5 and Fig. 15-6. Given the nature of the test, it seemed useful to divide the answers into three groups: perfect answers; answers which showed an arbitrary level of understanding of the concept required but which included incorrect calculations or other imperfections; and answers which merely showed that the pupil had attempted to solve the problem.

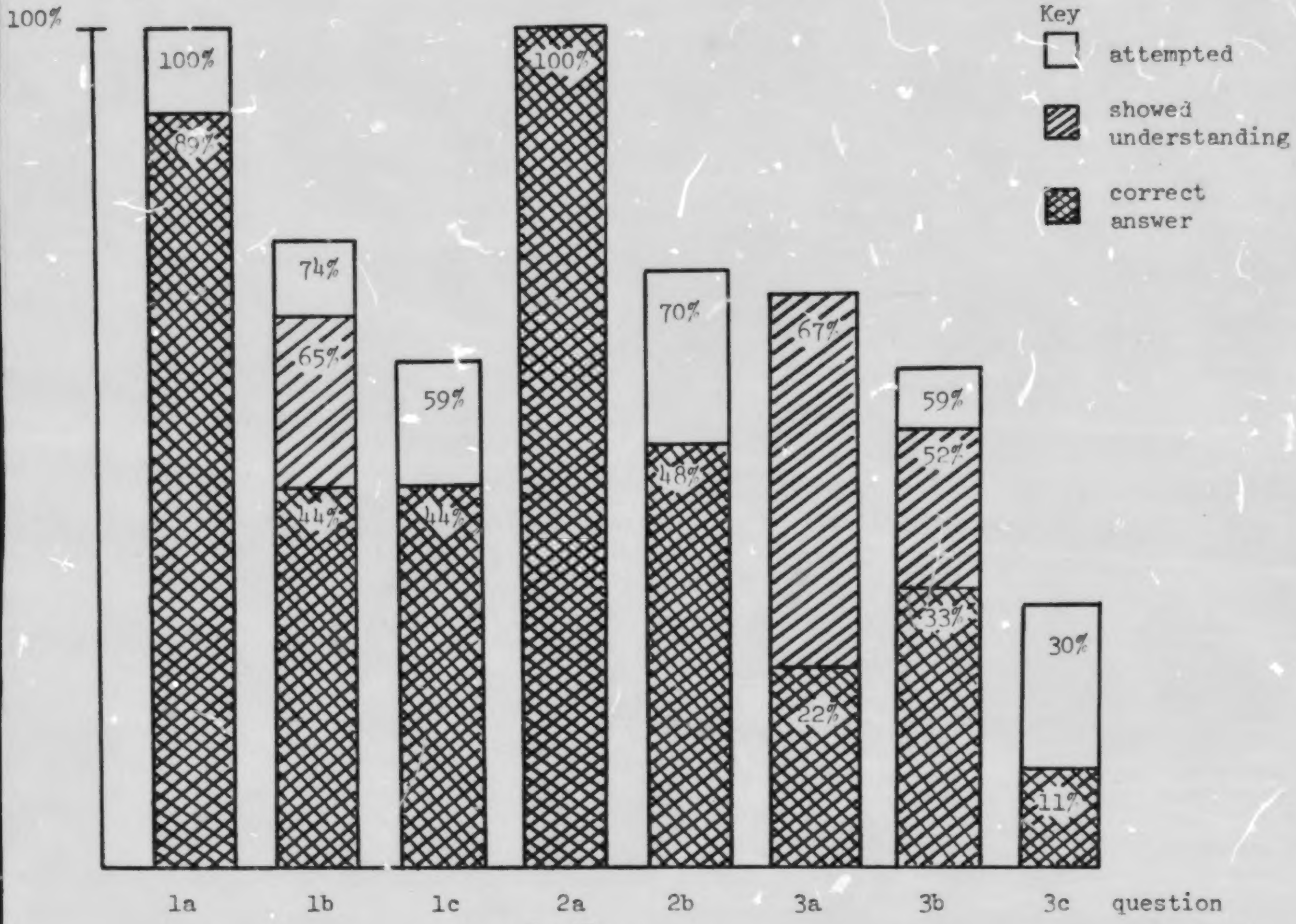
The instructions asked the student to answer the first question of each group (which was usually based on work done in class) and then to attempt the rest (which were often new to the student).

The problems caused by improper test construction are evident from the results in grade 5. For question 1, on each of parts a, b, and c, 9% of the students gave correct answers. Of the 20 students (65%) who attempted question 2a, 9 gave correct answers and 9 gave approximately correct solutions. On the other hand, not one student attempted part b. For question 3a there were 13 perfect answers out of 20 attempts, while for question 4 there was again not a single correct answer; the task was entirely new to the students.

The coverage of maths over a five year period was too great for us to do it justice within the time limit of our tests. It could be supposed that a longer testing period would have given significantly better results, especially in view of the time required for answering some of the questions.

Fig.15-5 Results of the modern mathematics tests in the Budapest study

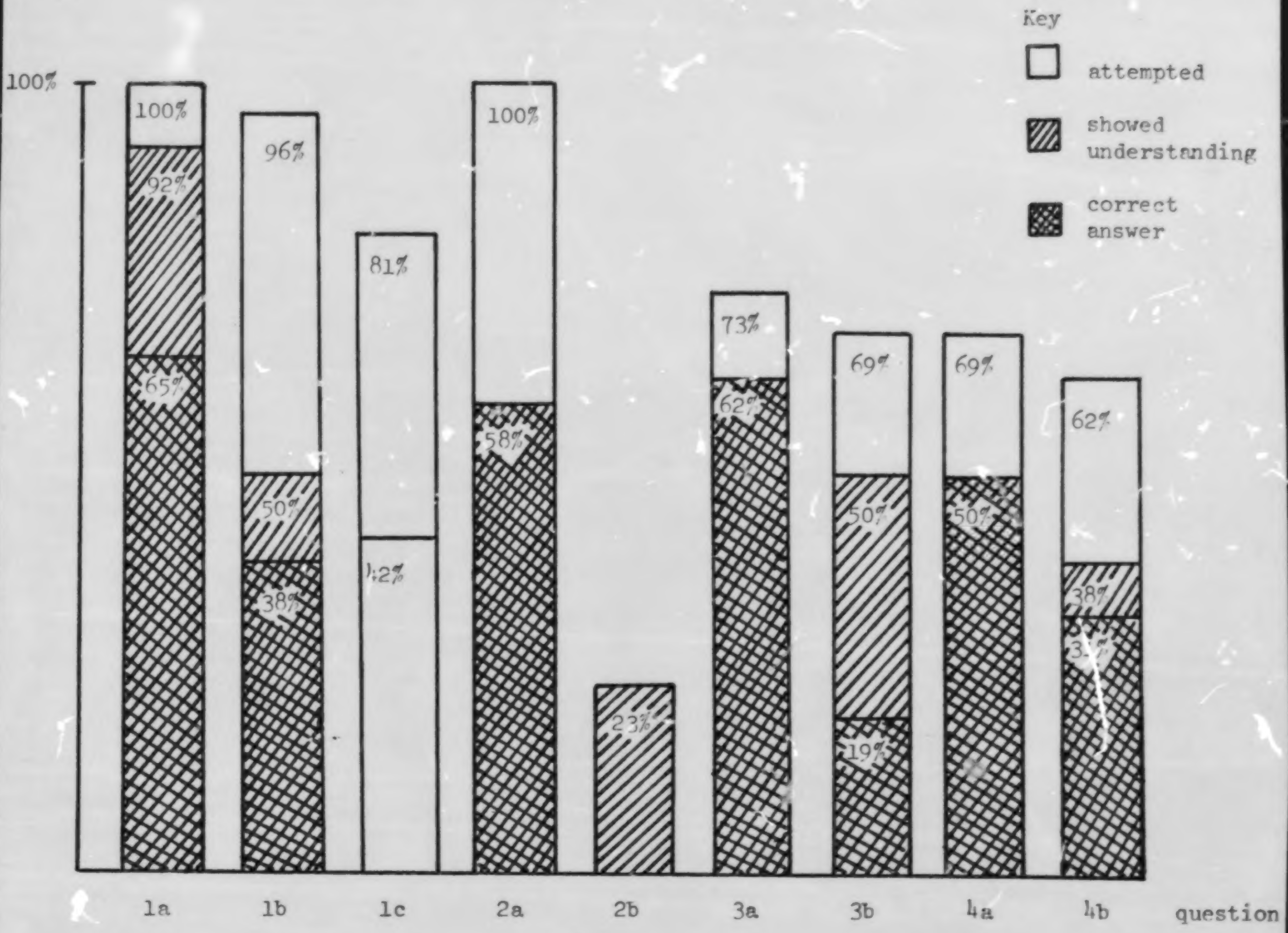
Grade 2 (N=27)



Notes: (1) Every student but one solved at least one subtask in question 1.
(2) In question 3, many students used numbers over 20.
(3) Already in grade 2 it seems natural forthe students to attempt problems with indefinite solutions.

Fig. 15-6 Results of the modern mathematics tests in the Budapest study

Grade 3 (N=26)



Note: (1) In question 3a, eight pupils gave 2-5 solutions.

In general, the results seem to support hypothesis 2. Children in the pilot classes seem to gain a great deal of knowledge of modern mathematics.

15.2.2 Sherbrooke Study

METHOD

Instrument: To examine the knowledge in modern mathematics of the Sherbrooke children, test booklets (see Appendix III) were constructed for grades 4, 5 and 6, each consisting of two parts.* Table 15-8 gives a list of the topics examined together with the number of questions on each for each grade.

Testing period: Knowledge in modern mathematics (as well as in traditional mathematics) was examined in May, 1973, about one year later than the other tests, the children belonging at the time of testing to grades 4, 5 and 6 respectively. The two parts of the test were given on successive days, or with only one day between.

RESULTS

The nature of the hypothesis allows us to present the data in a descriptive way only. As shown in the figures 15-7 to 15-9, children of the experimental classes are able to solve a certain number of problems in modern mathematics. However, the results varied very much from class to class in the different subtests. It seems that children of grade 6 (exp. group grade 5) have a better knowledge in this field than those of the lower grades, but we have to keep

---

* Selection and arrangement of test items was done by Bernard Héraud.

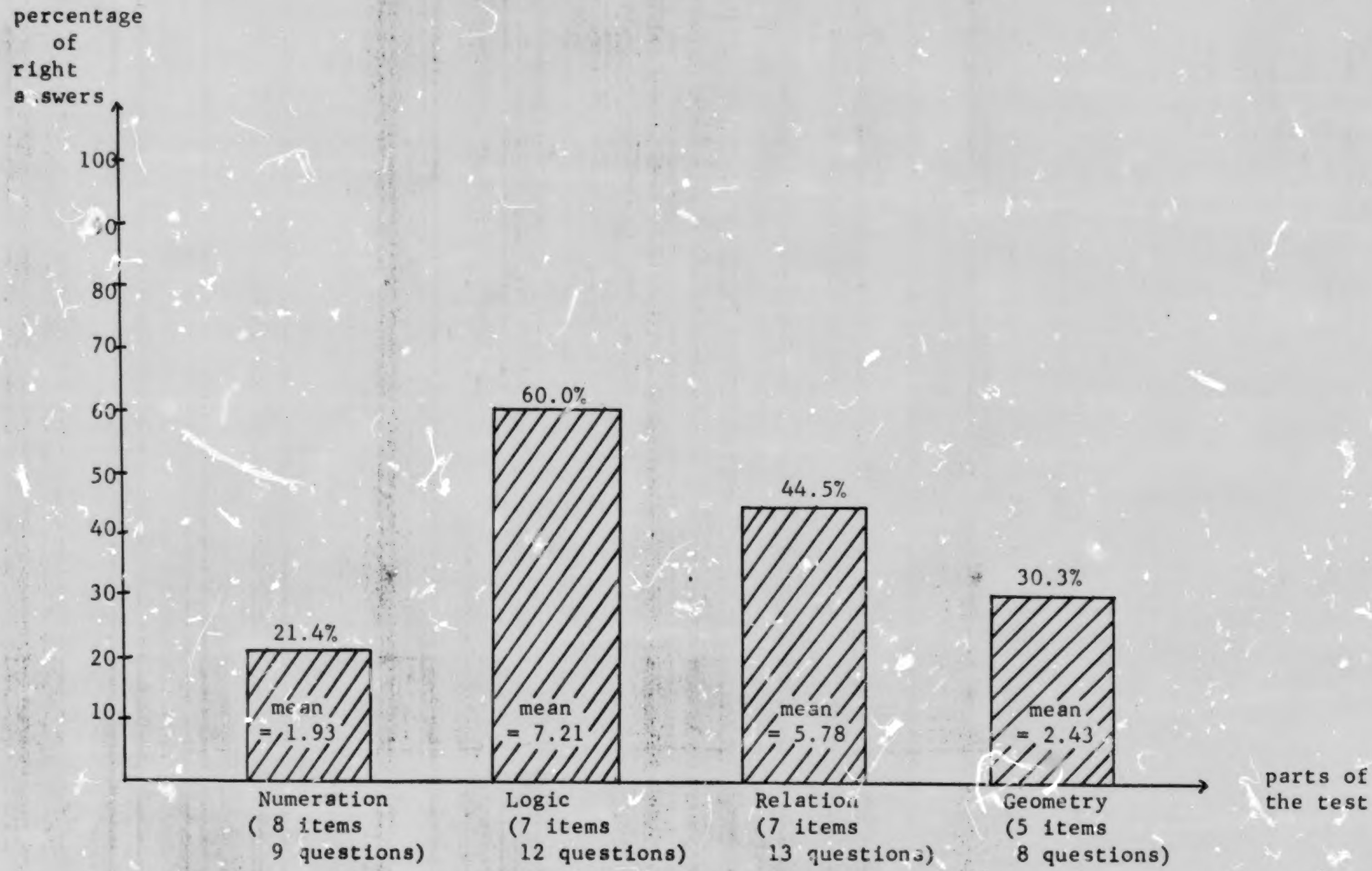


Fig 15-7

The results on different parts of the Modern Mathematics Test used in the Sherbrooke Study Grade 4 (identical with exp. group grade 3, N = 14)

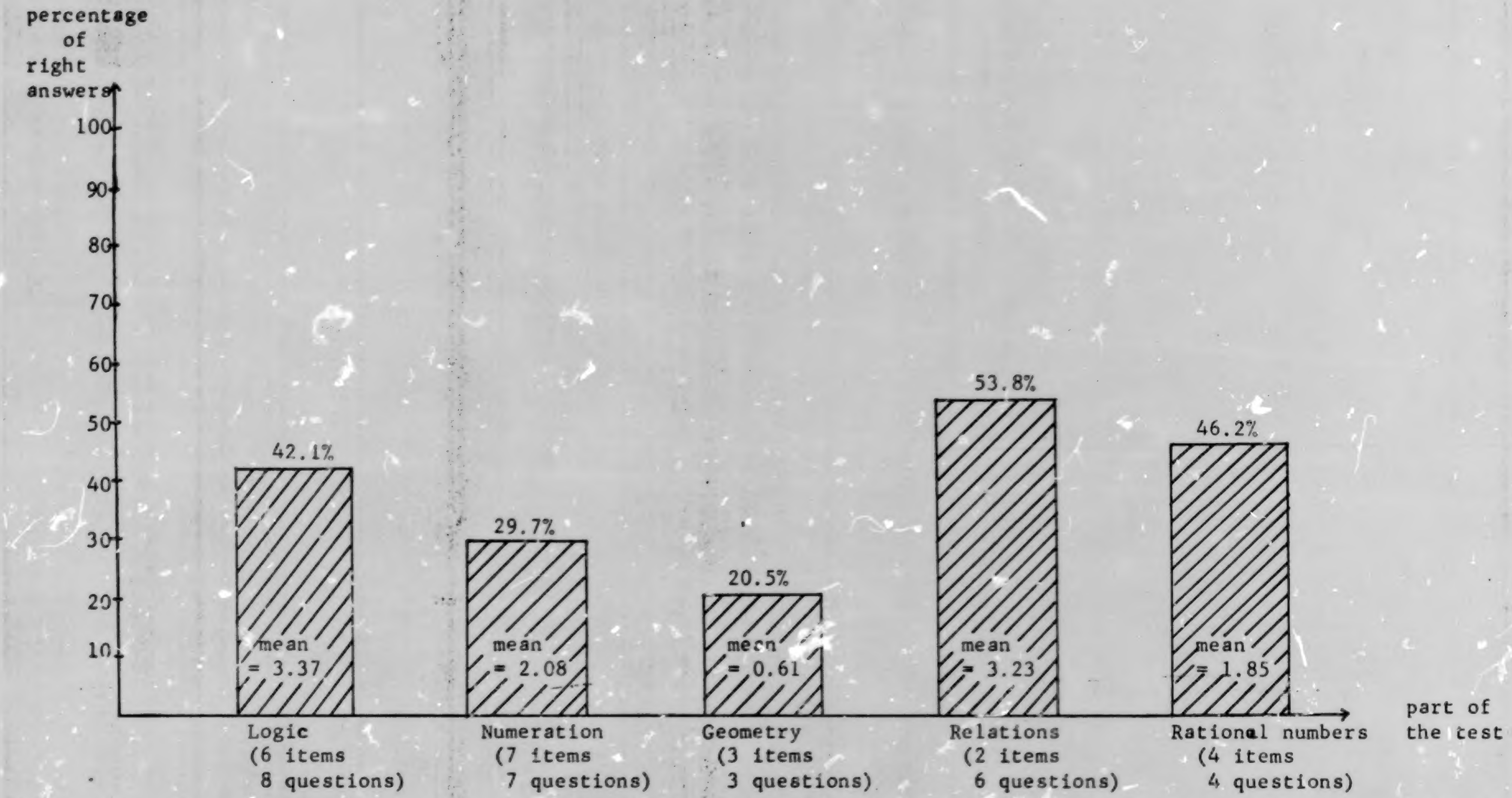


Fig 15-8

The results on different parts of the Modern Mathematics Test used in the Sherbrooke Study

Grade 5 (identical with exp. group grade 4, N = 13)

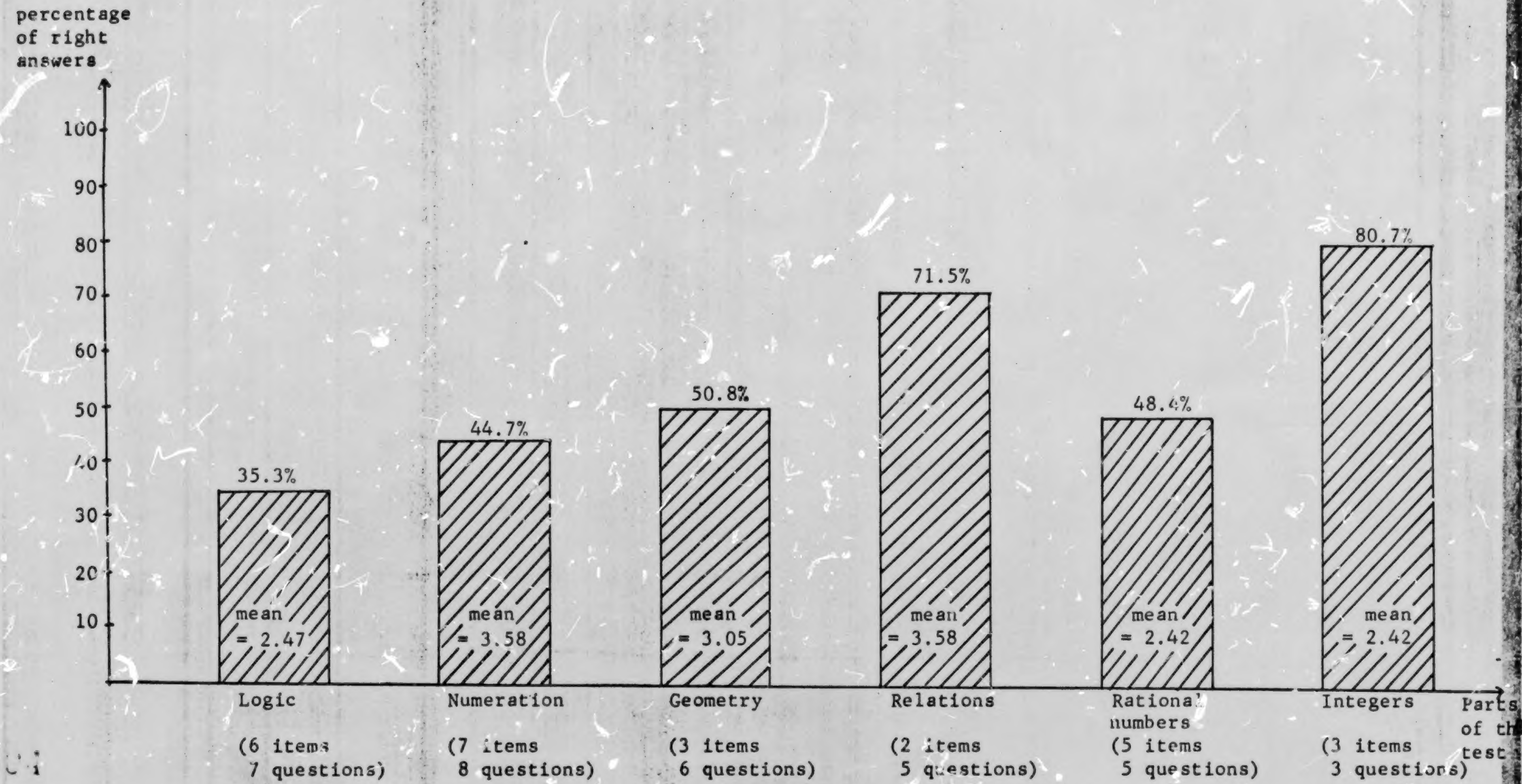


Fig 15-9

The results on different parts of the Modern Mathematics Test used in the Sherbrooke Study.

Grade 6 (identical with exp. group grade 5, N = 19) in mind that the results are based on "home-made-tests", which were, in regards of item difficulty and clarity of the text, not in a fully developed shape.[*] In consideration of this the results can be interpreted generally as showing a tendency towards our hypothesis that the children in the experimental classes gain a great deal of knowledge in the field of modern mathematics.

15.2.3 New York Study

The question of interest was: Is there any difference in the achievement in the field of "Modern Mathematics" between the students of Fleming school and those of Ecole Eymard?

METHOD

The same items which were chosen for the assessment of knowledge of the Eymard children at Sherbrooke were used here but 5 items (1 in grade 4 and 4 in grade 5) had to be left out from scoring because some mistakes slipped in and caused difficulties in comprehension.

The number of tasks on the different topics of mathematics examined are shown in Table 15-9.

TREATMENT OF DATA

To compare the two groups in each grade, the t-test was used on the results of each of the sub-tests as well as on the total.

---

[*] Concerning item difficulty we had to give to each item the same weight because the item bank is still in a developmental stage, and data which would allow for a differentiation of weights, not yet collected.

RESULTS

The results of the comparison between the Fleming School and the Ecole Eymard in the measurement of modern mathematics were not as clearly one-sided in favour of the first as they were in the case of traditional mathematics. In grade 4 Fleming School children were slightly superior only in the sub-test "Logic" (means of 8.29 against 7.21 based on 12 questions), and in grade 5 the same was true for the sub-test "Relations". The latter, however, consisted of only two questions and could thus not yield a very valid conclusion. The same was true of the sub-test "Rational Numbers" which contained only 2 questions, the Eymard children receiving the better score.

Taking all the questions together, there was no significant difference for children in either grade 4 or grade 5. In grade 5 the Fleming school children tended to be better (see Tables 15-10 and 15-11). Comparing the classes of the Fleming School with one another we found there was only a slightly better achievement by the "yellow" class of grade 4, and no remarkable difference in grade 5 (the difference in this grade had to be judged on the results of 21 questions only).

Table 15-5

The differences between the Fleming School in New York and the Eymard School in Sherbrooke in the IOWA Tests of Basic Skills Part: Mathematics Concepts.

Grades	School / Class / Statistics	Fleming School New York			Eymard School Sherbrooke
		red	yellow	total	total
4	mean	22.36	17.45	19.47	12.54
	stand. dev.	4.86	13.56	11.01	4.64
	No. of subj.	14	20	34	24
	F	—		F = 5.62	
	t			$t = -3.28^{**}$	
5	mean	26.29	23.00	24.75	15.90
	stand. dev.	8.65	13.08	10.89	4.74
	No. of subj.	17	15	32	21
	F	—		F = 5.27	
	t			$t = -4.04^{***}$	

*** significant at .001 percent level of confidence

** significant at .01 percent level of confidence

Table 15-6

The differences between the Fleming School in New York and the Eymard School in Sherbrooke in the IOWA Tests of Basic Skills Part: Mathematics Problem Solving

	School	Fleming School New York			Eymard School Sherbrooke
Grades	Class / Statistics	red	yellow	total	total
4	mean	17.21	19.40	18.50	12.75
	stand. dev.	4.10	6.43	5.62	5.97
	No. of subj.	14	20	34	24
	F t	———		F = 1.129 t =-3.70***	
5	mean	19.71	21.33	20.47	7.76
	stand. dev.	3.80	4.86	4.34	2.02
	No. of subj.	17	15	32	21
	F t	———		F = 4.60 t =-14.35***	

*** significant at the .001 percent level.

Table 15-7

The differences between the Fleming School in New York and the Eymard School in Sherbrooke in the IOWA Tests of Basic Skills Part: Mathematics Skills, (Problem Solving and Mathematics Concepts together).

Grades	School / Class / Statistics	Fleming School New York			Eymard School Sherbrooke
		red	yellow	total	total
4	mean	39.57	36.85	37.97	25.29
	stand. dev.	8.08	15.37	12.79	9.68
	No. of subj.	14	20	34	24
	F	—		F = 1.748	
	t			t = -4.29***	
5	mean	46.00	44.33	45.22	23.67
	stand. dev.	11.09	13.72	12.22	5.36
	No. of subj.	17	15	32	21
	F	—		F = 5.19	
	t			t = -8.77***	

*** significant at the .001 percent level of confidence.

Table 15-8

The number of questions on the different topics of mathematics examined with the Modern Mathematics Test in the Sherbrooke Study.

Part of the test	Topics	Number of questions examined in		
		Grade 4	Grade 5	Grade 6
1st part	Numeration (partly multibase)	9	7	8
	Logic	12	8	7
2nd part	Relations	13	6	5
	Geometry	8	3	6
	Rational Numbers	-	4	5
	Numeration (integers)	-	-	3

Table 15-9

The number of tasks and questions on the different topics of mathematics examined with the Modern Mathematics Test in the New York Study.

Part of the test	Topics	Number of questions scorable in		Number of tasks in	
		Grade 4	Grade 5	Grade 4	Grade 5
1st part	Numeration (partly multibase)	9	7	8	7
	Logic	12	8	7	6
2nd part	Relations	11	2	6	1
	Geometry	8	-	5	-
	Rational Numbers 1 + 2	-	4	-	4

Table 15-10

The differences between the students of the Fleming School (New York) and the Eymard School (Sherbrooke) in the achievement test of Modern Mathematics: Grade 4

Parts of the tests		Fleming School red N = 17 $\bar{X}$	Fleming School red N = 17 st. dev.	Fleming School yellow N = 18 $\bar{X}$	Fleming School yellow N = 18 st. dev.	Fleming School total N = 35 $\bar{X}$	Fleming School total N = 35 st. dev.	Eymard School total N = 14 $\bar{X}$	Eymard School total N = 14 st. dev.	F	t
1st part	Numeration (9 questions)	1.24	1.35	1.78	1.52	1.51	1.44	1.93	1.82	1.58	0.76
	Logic (12 questions)	7.71	1.49	8.67	1.71	8.20	1.66	7.21	1.48	1.26	-2.03*
2nd part	Relations (11 questions)	4.47	2.92	6.33	1.61	5.43	2.49	5.07	2.59	1.08	-0.44
	Geometry (8 questions)	1.41	1.42	1.78	1.40	1.60	1.40	2.43	1.28	1.18	1.98
Total (40 questions)		16.82	5.74	18.56	4.10	16.74	5.24	16.64	5.67	1.16	-0.05

* significant at the .05 percent level of confidence.

Table 15-11

The differences between the students of the Fleming School (New York) and the Eymard School (Sherbrooke) in the achievement test of Modern Mathematics: Grade 5.

Groups / Parts of the test / Statistics		Fleming School						Eymard School			
		Orange N = 15		Yellow N = 13		Total N = 28		Total N = 13			
		$\bar{X}$	st. dev	$\bar{X}$	st. dev.	$\bar{X}$	st. dev	$\bar{X}$	st. dev.	F	t
1st part	Logic (8 questions)	4.13	1.46	4.54	1.66	4.32	1.54	3.77	1.74	1.27	-0.98
	Numeration (7 questions)	2.93	1.39	3.15	1.63	3.04	1.48	2.08	1.61	1.18	-1.82
2nd part	Relations (2 questions)	1.27	0.70	1.54	0.66	1.39	0.69	0.23	0.60	1.30	-5.51***
	Rational numbers$_1$ (2 questions)	0.73	0.80	0.77	0.83	0.75	0.80	0.77	0.73	1.21	0.07
	Rational numbers$_2$ (2 questions)	0.53	0.64	0.54	0.78	0.54	0.69	1.08	0.76	1.53	2.18*
Total (21 questions)		9.60	2.95	10.54	3.86	10.04	3.37	7.92	3.40	1.01	-1.85

*** , significant at the .001 percent level of confidence.

* significant at the .05 percent level of confidence.

Chapter 16

Fatigue

Hypothesis 3: The children in the pilot classes are not more tired after their mathematics lesson than the children learning arithmetic in the traditional way.

The third hypothesis* was concerned with student fatigue caused by the mathematics lesson. Unfortunately, there is no single method available which can be a reliable measure of the fatigue effect of such a short mental activity (Lehman, 1953; McFarland, 1971).

The introspective methods of measuring mental fatigue have shown great development in the last few years (e.g. the fatigue rating score designed by Kashiwagi, 1971), but they are much more useful with adults than with children. The physiological and psycho-physiological parameters are affected much more by physical work than by mental work but in appropriate combinations they are also useful for measuring mental fatigue (Róna et al., 1971).

16.1 Treatment of data

To examine the third hypothesis we calculated the differences between the test results obtained before and after the mathematics lesson and used the one sample t-test to determine their significance. We considered significant differences as a sign of fatigue. If there were no significant differences - either in the experimental or control group - then we concluded that the particular parameter in the particular age group was not sensitive enough to show the fatigue caused by only one mathematics lesson. If a significant difference was noted in either group, then the changes caused by the mathematics lesson were compared using the following methods: A two-sample t-test if the F ratio showed no significant difference between the standard deviations of the measurement changes; and a d-test (see Hajtman, 1969), if it did.

---

* The variable 'fatigue' was only examined in the Budapest Study.

16.2 Theoretical background and methods

The measurements of fatigue, used in grades 2 through 5, were the following:

16.2.1 Reaction time for visual stimulus

Reaction time is probably the most widely used measurement of both mental and physical fatigue, because fatigue causes a quite noticeable lengthening of it. Its use is quite common even in educational situations (e.g. using this parameter Geréb, 1962, could differentiate between the fatigue effects caused by morning and afternoon school sessions). The apparatus we used in our experiment consisted of three parts (Fig. 16-1): a generator emitting a light or sound stimulus, a response button, and a timer measuring the time between stimulus and response. The instrument gave each subject five visual stimuli. For each person, we took into consideration the two smallest times larger than 130 msec. (the minimum permitted by the instrument). The average of the two reaction times was considered the subject's score.

Expectation: Equal or smaller increment in visual reaction time in the experimental classes than in the control classes after the mathematics lesson would show equal or less mental and physical fatigue, thus supporting Hypothesis 3.

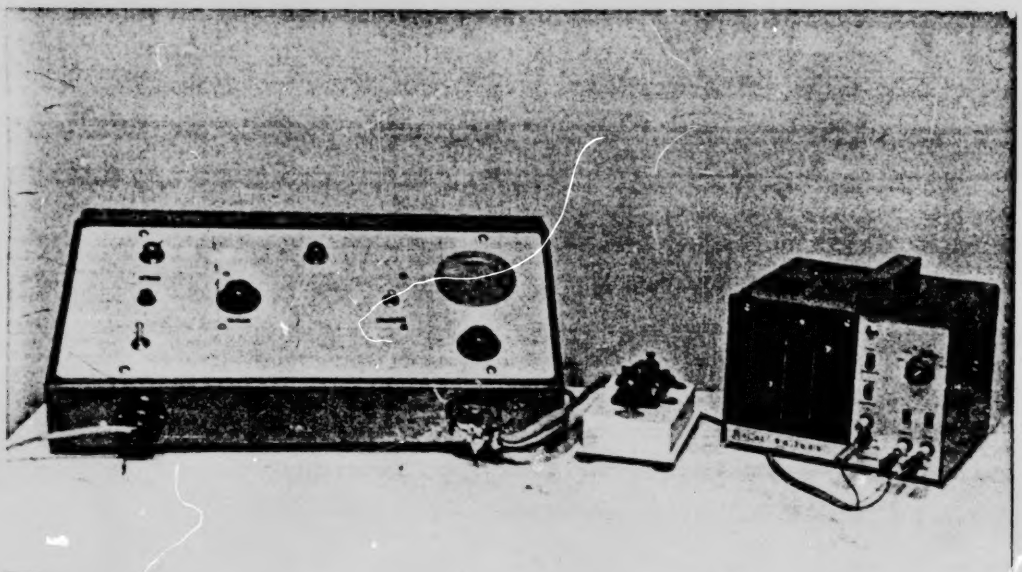


Fig. 16 - 1

Instrument for measuring reaction time.

16.2.2 Reaction time for acoustic stimulus

The apparatus, the process and the expectation were the same as in 16.2.1.

16.2.3 Heart rate

It is well demonstrated especially in industrial and sports psychology that heart rate is higher after physical effort. On the other hand, mental fatigue causes a drop in heart rate. Róna et al., (1971), for example, found a significantly bigger decrement after a school day in a group of children (aged about 10) who learned in worse school conditions, than in those whose learning situations were better. The apparatus we used to measure heart rate consisted of two parts (Fig. 16-2), a transmitter and a receiver. The transmitter registered the impulses from the student's finger. The receiver was connected to a counter and the heart rate could be read off directly in beats per minute.

Expectation: It is difficult to interpret the results of this test, as mental and physical fatigue cause opposite changes, possibly masking one another. Nevertheless, if we assume that a mathematics lesson causes minimal physical fatigue in comparison with mental fatigue, then we can say that equal or smaller decrement of heart rate in the experimental classes than in the control classes after the mathematics lesson would show equal or less mental fatigue, thus supporting hypothesis 3.

16.2.4 Visual field

Generally, the efficiency of the sense organs diminishes with fatigue.

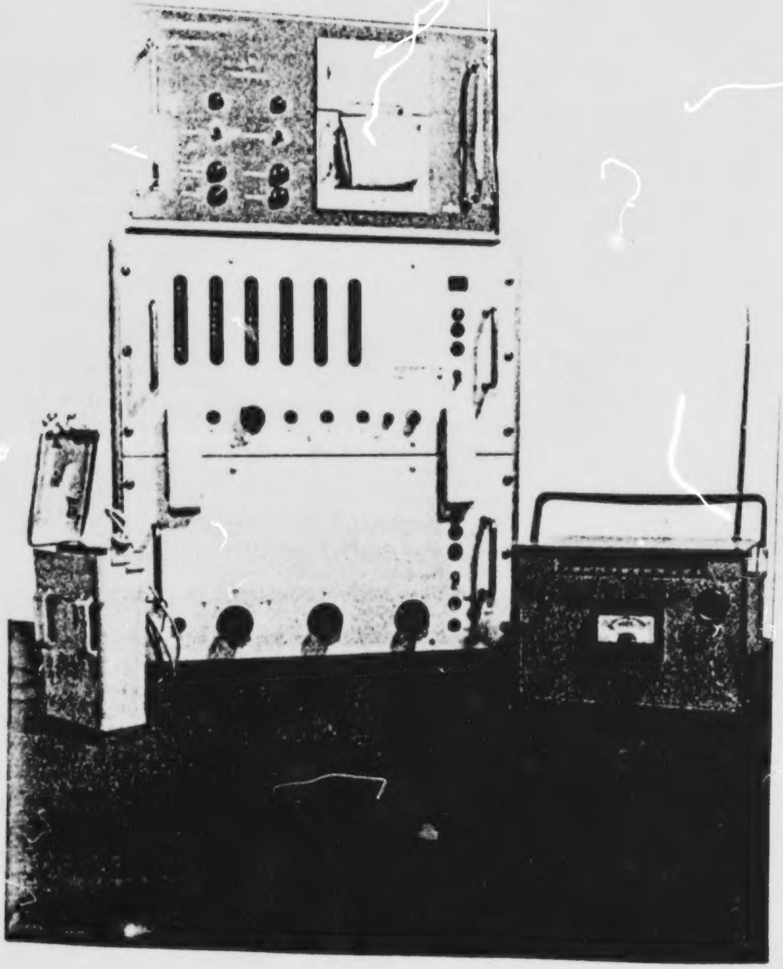


Fig. 16 - 2

Instrument for measuring heart rate

Unfortunately, atypical reactions make this a difficult effect to demonstrate. Horváth (1963) for example found an increment in 7 to 10 percent of a sample of truck drivers after 10 to 11 hours of work. He also found some indication that this atypical effect occurs more often in the worker whose attitude toward his job is favourable. The apparatus we used to measure the students' visual field was the hemispherical perimeter (Kugel-perimeter: Carl Zeiss Jena; Fig. 16-3). To shorten procedure we measured only the horizontal visual angle. The light spot was white, its surface was 64 mm^2 and its luminous flux was 1500 apostilbs. (The luminous flux of the inner surface of the hemisphere on which the light spot moved was 50 apostilbs.)

For each trial with each student we got four data: two when the light spot was moving from the edge of the hemisphere to the middle and the subject perceived its appearance, and two when the light spot was moving from the middle of the hemisphere to the edge of the hemisphere and the subject perceived its disappearance. We took into consideration the average of these four dots and so, as with the reaction time we got two figures for each student, one before and one after the mathematics lesson.

Expectation: Equal or smaller decrement in the visual field in the experimental classes than in the control classes after the mathematics lesson would show equal or less fatigue thus supporting hypothesis 3.

16.2.5 Finger tremor

Tremor is the unintentional oscillation of a certain part of the body. The phenomenon itself is very complex: according to Wyatt (1968), at least 37 parameters have an effect on it. Because of this confusion, tremor alone

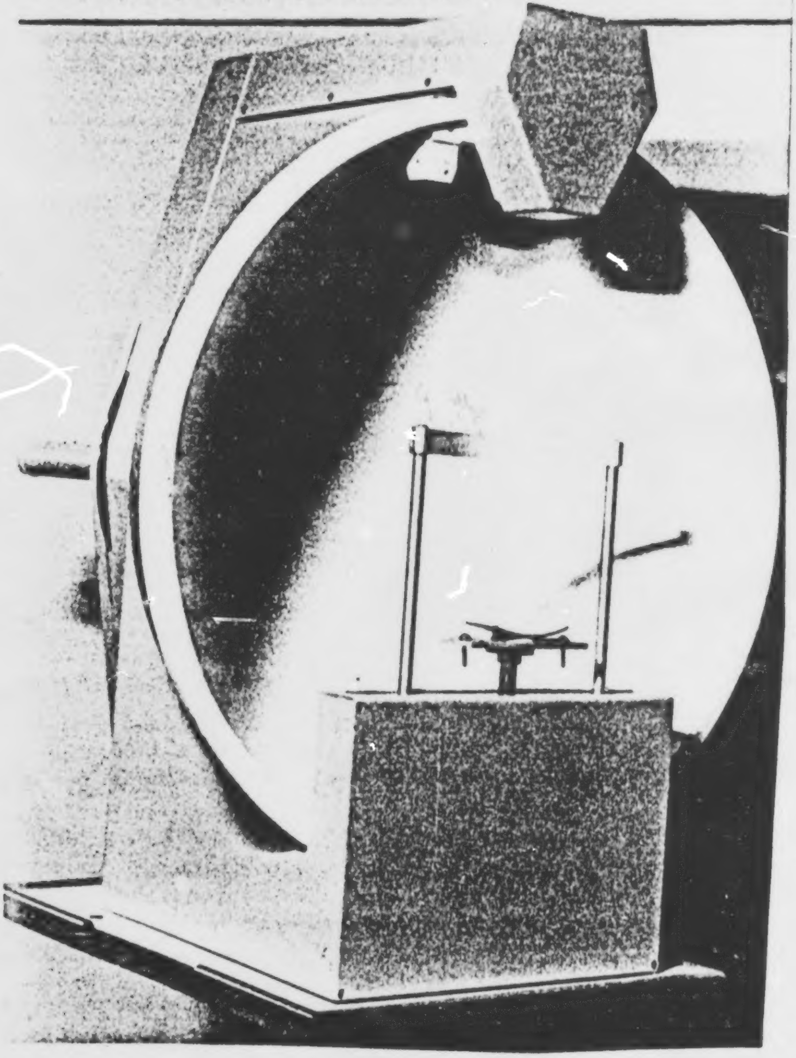


Fig. 16 - 3

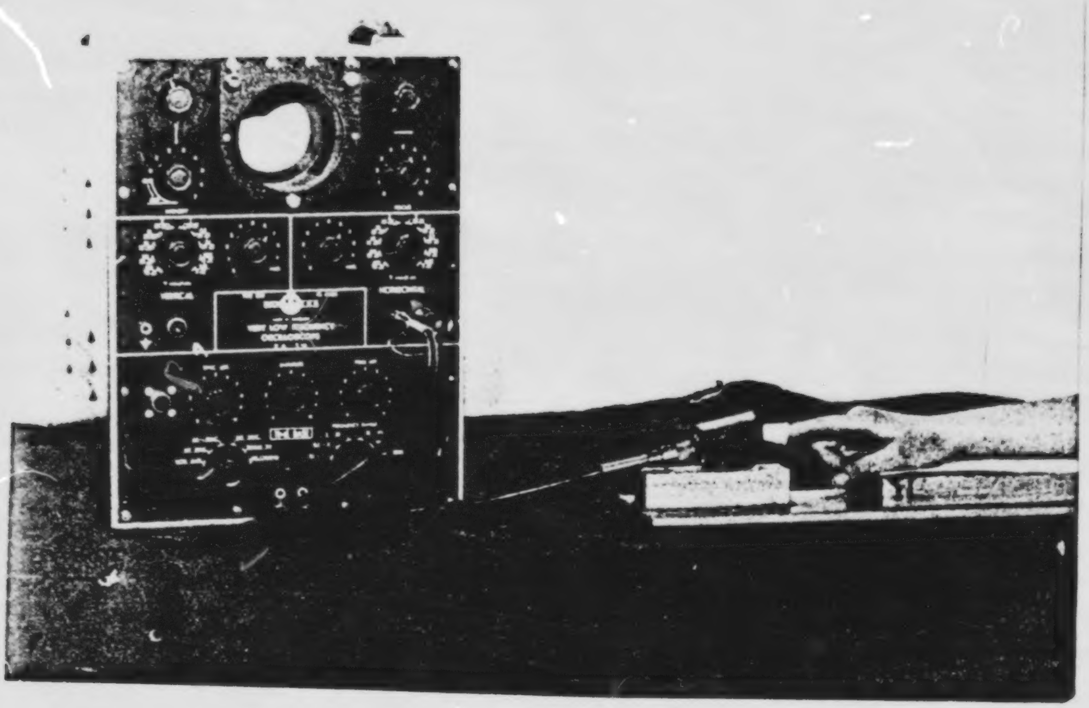
Hemispherical perimeter (Kugel-perimeter; Carl Zeiss Jena) for measuring visual field.

is not an effective way to measure fatigue. Together with other measurements, though, it is a useful tool, especially to differentiate between mental and physical fatigue: Under physical fatigue the amplitude of the vibration decreases while under mental fatigue the frequency of the vibration increases. (For further discussion of tremor, see for example Haliiday and Redfern, 1956; and Brumlik, 1962.) There is quite a variety of instruments measuring tremor. The essence of our instrument (Fig. 16-4) was a small permanent magnet fixed in a thimble which was placed on the subject's forefinger. This permanent magnet moving before an electro-magnet induced a current which was transmitted to an oscilloscope and recorded by a photo-recorder. The distance between permanent and electro-magnets was kept carefully constant. The instrument gave reliable information only about the frequency of the tremor.

Expectation: Equal or smaller increment of finger tremor frequency in experimental classes than in control classes after the mathematics lesson would show equal or less mental fatigue thus supporting hypothesis 3.

16.3 Results

The results of the examination of student fatigue caused by mathematics lessons are summarized in Table 16-1. As can be seen in the table, of the 34 possible differences concerning the fatigue caused by the mathematics lesson, 24 were considered significant, thus showing that our measurements were quite appropriate to the problem. It is unfortunate that because of a technical problem with the instrument, scorable results for finger tremor were obtained



VERTICAL
HORIZONTAL
VERY LOW FREQUENCY
OSCILLOSCOPE

Fig. 16 - 4

Instrument for measuring finger tremor.

Table 16-1 Summary of the results of the examinations of the students' fatigue.

Class		Reaction time for visual stimulus	Reaction time for acoustic stimulus	Heart rate	Visual field	Finger tremor	Number of significant differences
2	Experience	*	*	*	0	0	3
	Control	0	*	0	*	0	2
	Difference	$\boxed{0}$	$\boxed{0}$	+	$\boxed{0}$	■	1
3	Experimental	0	*	0	0		1
	Control	0	*	*	0		2
	Difference	■	$\boxed{0}$	-	□		1-
4	Experimental	*	*	0	*		3
	Control	*	*	0	0		2
	Difference	$\boxed{0}$	$\boxed{0}$	□	$\boxed{0}$		0
5	Experimental	*	*	0	0		2
	Control	0	*	0	0		1
	Difference	+	+	■	■		2
Number of significant differences showing fatigue		4	8	2	2	0	16
Number of significant differences between the experimental and control classes		1	1	0	1+1-	0	3+1-

* significant difference in the parameter after mathematics lesson ($p < 0.05$)
0 no significant difference in the parameter after the mathematics lesson ($p > 0.05$)
+ the changes in the parameter show a significantly greater fatigue in control class ($p < 0.05$)
- the changes in the parameter show a significantly greater fatigue in experimental class ($p < 0.05$)
$\boxed{0}$ no significant difference in the changes between experimental and control class ($p > 0.05$)
■ the parameter was inappropriate for measuring fatigue in this class.

in grade 2 only.

As the table shows as well, very few significant differences were found between the experimental and control classes: only 4 out of the 17 possible. Three of these differences (heart rate in grade 2 and reaction time for visual and acoustic stimuli in grade 5) were in the direction of greater fatigue in the control classes, while one (heart rate in grade 3) was in the opposite direction. In general these results support our hypothesis.

16.3.1. Reaction time for visual stimulus

The mean reaction time for the six groups ranged between 248 and 355 msec. The mathematics lesson caused an increment of between 1 and 47 msec. in the average reaction times of the groups, a result which was significant in grades 2 ($t = 4.533$); 4 ($t = 3.349$); and 5 ($t = 3.661$) of the control group and in grade 4 of the experimental group ($t = 3.349$). We found a significant difference between the experimental and control groups only in grade 5 ($t = 2.800$); the lesser fatigue effect being in the experimental class. Because we found significant differences neither in the experimental nor in the control group in grade 3, comparisons were not made in this grade. The t-test ($t = 1.690$) in grade 4 and the d-test ($d = 1.939$) in grade 2 showed no significant differences.

16.3.2 Reaction time for acoustic stimulus

The mean reaction times ranged between 222.2 (in grade 5 before the mathematics lesson) and 343.3 msec. (in grade 2 after the mathematics lesson). The mathematics lesson caused mean increments of between 15 and 77 msec., which were significant in all cases. Because with all the other parameters significant results were obtained only in some cases, reaction time proved to be the most sensitive measurement of the fatigue caused by the mathematics lesson. Because the F-ratios showed no significant differences between the SD's of the results in experimental and control groups, we could use t-tests in each case; they showed only one significant difference: a greater increment in the grade 5 control class (t = 2.888).

16.3.3 Heart rate

The group averages of student heart rate varied between 79 and 101 beats per minute. The mathematics lesson caused a significant drop in only two groups: 2 beats per minute in the grade 2 control group (t = 5.340) and 6 beats per minute in the grade 3 experimental group (t = 4.117). In both grades, the difference between the experimental and control group was significant (t = 3.606 in grade 2 in favour of the experimental group and d = 2.793 in grade 3 in favour of the control group.)

16.3.4 Visual field

The average visual angles were between 61° and 82°. The mathematics lesson caused a significant drop in only the grade 2 experimental (2°; t = 2.580) and grade 4 control (5°; t = 4.587) groups. But while in the latter case there was a significant difference between the experimental and control groups (t = 2.165), the difference was not significant in grade 2 (t = 0.199).

16.3.5 Finger tremor

Because of trouble with the instrument we got appraisable results only in grade 2. But even these results did not show significant changes as an effect of the mathematics lesson.

Chapter 17

Attention

Hypothesis 4: The mathematics lesson in the Dienes classes causes a greater concentration of the children's attention than those in the control classes.

17.1 Theoretical background

The concept of attention* has a short scientific history. Although "attention" was the cornerstone of consciousness-centered psychologies at the turn of the century (Pillsbury, 1908; Titchener, 1908; Wundt, 1897), the many problems connected with it were not solved and there was practically an absence of continued development of research and theory in this field for many years thereafter. Only since the second world war have psychologists - many from the Soviet Union (see e.g. Simon, 1957) - become interested in individual differences in higher nervous activity and invested new energy in an examination of its characteristics, having only begun to realize its importance in the fields of mental health and especially in education (Day, 1967). At this time, we are still beginning to discover the main principles connected with it (Maltzman, 1967).

An accepted truism states that learning cannot take place unless attention is attracted and maintained in the learning situation, or especially in our case when the student's activity during mathematics lessons, both in traditional and experimental classes, requires a certain concentration of attention. Logically, attraction of students' attention towards highly complex material -

---

* The variable 'attention' was examined only in the Budapest Study.

like that often used in mathematics lessons - is strongly related to the arousal level within the students (Berlyne, 1960; Berlyne and Lewis, 1963; Day and Llewellyn, 1967): For example, new, and complex materials would attract students' attention and so induce a higher state of arousal. In fact, the results of several studies in this field are not so simple: While some papers even report a decrease in exploration under conditions of arousal (Berlyne and Lewis, 1963; Haywood, 1962; McReynolds and Bryan, 1956; Smock, 1955a, b), others fail to find any change (Day, 1965; Driscoll and Lanzetta, 1964; Hawkins and Lanzetta, 1965; Haywood, 1961; Haywood and Hunt, 1963). Only very few studies (e.g. Day and Llewellyn, 1967) have found an increase in attention for high complex material with an increase in the level of arousal.

Accepting Hebb's (1955) postulation of an inverted U-shaped distribution of affect as well as an effective "cue function" over arousal, we should expect that the proportion of attention towards highly complex material would rise to a peak with increased arousal and then drop off as arousal level continues to rise (Day, 1967). In our case, we supposed that the students entering school in the morning (the mathematics lesson was always first), could during their mathematics lesson be placed upon the first half of this inverted U-shaped distribution; we continued to suppose that an increase in the level of arousal after the mathematics lesson meant a greater amount of attention paid to the preceding mathematics lesson.

17.2 Method

There are some indications that achievement in a previously learned and practised monotonous activity depends more on the degree of attention accorded to the preceding activity (or on the level of arousal reached) than on the fatigue caused by it (Geréb, 1970b). An interesting, intriguing activity increases achievement on the following monotonous activity. In the light of this fact, to examine our fourth hypothesis we chose the Toulouse-Pieron test (see e.g. Pieron, 1929), which consists of a typical monotonous task. In this test, there are 20 "eared" squares on a page. The "ears" are on one of the sides or the corners of the square. The task is to find usually 4, in our case, 2, particular types of eared squares. To measure the efficiency of the subjects, we used the following formula:

$$T = \frac{(N - H) \times 100}{N} \qquad \text{Eq. 17-1}$$

where N = the number of appropriate squares found by the subject; and

H = the number of errors made by the subject.

Although the test is widely used in different fields of applied psychology, some researchers strongly criticize it (e.g. Kanizsa, 1951, stating that its diagnostic value is slight; it measures different variables at different subject levels, the qualitative and quantitative parameters - N and H - in T are not correlated; etc.). We, however, felt that it could serve our special purpose. To eliminate the effect of practice, we allowed the children to practice the task at length during preceding days. The tests of the fourth hypothesis were carried out in grades 2 through 5.

17.3 Treatment of data

The statistical treatment of the fourth hypothesis was the same as that of the third.

17.4 Expectation

Larger score increments in the experimental classes than in the control classes on the Toulouse-Pieron test performed immediately after the mathematics lesson would show greater attention paid to the preceding activity, thus supporting hypothesis 4.

17.5 Results

The results of the Toulouse-Pieron test showed a significant increment in student efficiency in performing this monotonous task in all grades. While the average scores showing this efficiency were between 36 and 76 before the mathematics lesson, they were between 47 and 90 after the mathematics lesson with an increment ranging between 11 and 25. (The results of the t-tests were between t = 3.200 and t = 10.404.) However, we did not find any significant differences between the experimental and control group increments (the t-tests gave results of 0.770 and 1.094 in grades 3 and 4, while the d-test gave results of 1.428 and 0.360 in grades 2 and 5). So while this measurement - as the significant differences caused by the mathematics lesson showed - is a sensitive one, it still failed to show the difference we expected to obtain in our hypothesis.

17.5.1 Discussion of the results

There are several possible explanations of this hypothesis. Besides the simple explanation that the mathematics lesson may be as interesting for the students in the control classes as in the experimental classes, there exist other equally valid possibilities. For example, the cause of the Toulouse-Pieron test result differences after a school lesson is probably due to an increment in the students' arousal. But this can be the result either of an internal interest (unintentional attention on work or play, etc.) or an external pressure (intentional attention on work, avoidance of recitation, etc.). The real drawback of this method is that it fails to differentiate between these two possibilities.

We wish to point out here another difficulty we did not take into consideration at the time we designed this evaluation. As Day (1967) suggests in his explanation of the data and in modifying Hebb's model, the arousal-attention function like student performance in learning situations (Spence, 1956) is different in high-anxious and low-anxious students (see Fig. 17-1). While attention is proportional of the level of arousal in low-anxious children, it is inversely proportional to the level of arousal in high-anxious children. Unhappily, we did not obtain our results separately in the different anxiety groups, but if Day's model is true, then of course, no consequence could be drawn from our data even if we were to find significant experimental-control differences. A higher rise in the level of arousal after the mathematics lesson in experimental classes could indicate higher student attention in low-anxious students and lower student attention in high-anxious students. So, as an average, it could indicate the same level of attention.

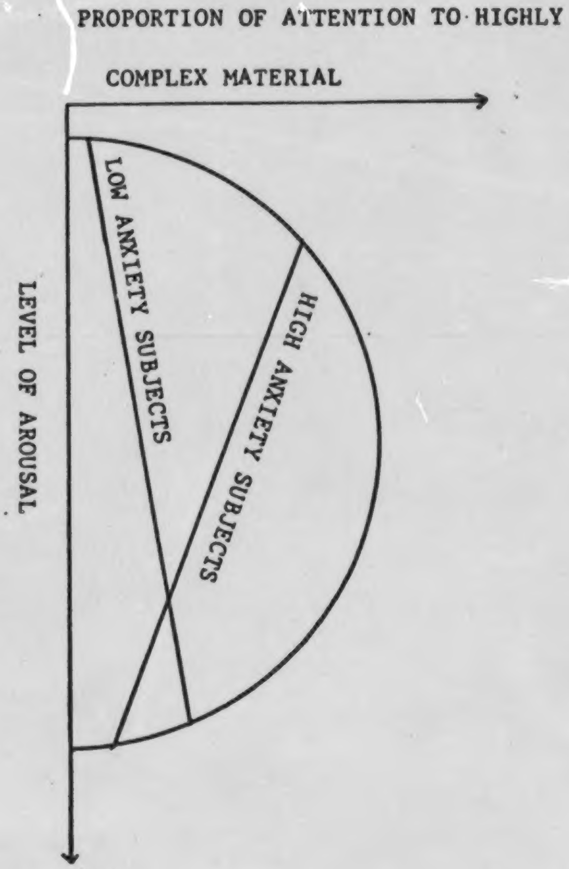

PROPORTION OF ATTENTION TO HIGHLY COMPLEX MATERIAL
LEVEL OF AROUSAL
LOW ANXIETY SUBJECTS
HIGH ANXIETY SUBJECTS

Fig.17-1 The anxiety-attention function (after Day, 1967).

Chapter 18

Learning ability

Hypothesis 5: The Dienes method of mathematics learning helps to develop the students' learning ability more than the traditional methods.

It seems to be a reasonable hypothesis that the Dienes mathematics teaching method with its student-centered and learning-centered arrangements, in which so many learning decisions are made by the students, helps to develop the students' learning ability, especially their ability to learn material related to mathematics. However, it has been rather difficult to find a method of measuring the children's learning ability.

An attempt to find such a measurement began with the development of intelligence tests: from Binet and Simon (1905) to Wechsler (1949), it appears very commonly in definitions of intelligence that it is "the ability to learn", the "ability to adapt to new situations", etc. But in fact the traditional intelligence tests have approached the very dynamic concept of learning ability in a very static way. Therefore we decided to construct a new test: The Secret Writing Test.

This test is designed to measure the learning ability of students through a real learning process, providing separate and combined information about the speed and the effectiveness of the learning process and about the child's ability to creatively apply his new knowledge. The theoretical background and the method of measuring learning ability will be described later in more detail.

18.1 Theoretical background

18.1.1 The Origin of the idea

Apart from some early pioneering studies by Woodrow and others (see section 18.1.2), psychologists, especially experimental psychologists, began to be interested in individual differences in human learning only in the last fifteen years (Fleishman, 1967). However, the idea of our test goes back to the early 1930's. L.S. Vygotsky, the Russian psychologist, proposed at that time a new concept of intelligence (as a capacity to benefit from instruction) together with a new concept of intelligence testing. In his book, Thought and Language, he wrote:

> Most of the psychological investigation concerned with school learning measured the level of mental development of the child by making him solve certain standardized problems. The problem he was able to solve by himself was supposed to indicate the level of his mental development at the particular time. But in this way only the completed part of the child's development can be measured, which is far from the whole story. We tried a different approach. Having found that the mental age of two children was, let us say, eight, we gave each of them harder problems that he could manage on his own and provided some slight assistance, the first step in a solution, a leading question, or some other form of help. We discovered that one child could, in co-operation, solve problems designed for twelve-year-olds, while the other could not go beyond problems intended for nine-year-olds. The discrepancy between a child's actual mental age and the level he reaches in solving problems with assistance indicates the zone of his proximal development; in our example, this zone is four for the first child and one for the second. Can we truly say that their mental development is the same? Experience has shown that

> the child with the larger zone of proximal development will do much better in school. This measure gives a more helpful clue than mental age does to the dynamics of intellectual progress (1962).

Because of the language barrier, however, Vygotsky's ideas (with some notable exceptions; e.g. Haufmann and Kasanin, 1937) did not have their due effect on the development of mental testing until recently.

18.1.2 Studies with simple learning tasks in the U.S.A.

Independently of Vygotsky's work, the possibility of measuring learning ability by using simple learning tasks was examined by several researchers (e.g. Husband, 1939; 1941a, b; Woodrow, 1938; 1939a, b, c; Heese, 1942; Carlson et al., 1945; Thompson and Witryol, 1946) in the late 30's and 40's. Their work is well summarized in Guilford's book, The Nature of Human Intelligence (1967a).

Three kinds of measurements of learning ability were tried out:

1. Learning score; that is the mean performance in learning tasks. The main problem with this is that unless all persons begin at the same level it is clearly unsatisfactory.

2. Gain score; that is the difference between final and initial scores in a task or some function of that difference. The main problem with using this measurement is that gain scores are usually very unreliable (see Gulliksen, 1967). According to Guilford (1954), the reliability of the gain score (r_{gg}) is calculated by the following equation:

$$r_{gg} = \frac{r_{ff} + r_{ii} - 2r_{fi}}{2(1 + r_{fi})} \qquad \text{Eq. 18-1}$$

where r_{ff} = the reliability of the final test,

r_{ii} = the reliability of the initial test; and

r_{fi} = the correlation between the initial and the final test.

From this equation, it can be determined that unless the correlation between the initial test and the final test is very low (which can be reached only through the use of a very long practice period), the reliability of the gain score will be low even if the reliability of both the initial and final tests are high. (For a discussion of the several difficulties in using and interpreting gain scores, see the excellent book on measuring change edited by Harris, 1963.)

3. Slope-rate score; that is a certain parameter of the curve fitted to the cumulative distribution of the practice score of the subject. The main problem with this unquestionably most sophisticated method is in defining what kind of function to fit to the empirical data. The final decision is usually an arbitrary one as can be demonstrated by the variety of functions used and the parameters thought to be necessary to describe the learning record: Cumulative normal distribution curves (Thurstone, 1930; Guilford, 1967a), the hyperbola with one (Wiley and Wiley, 1937), with two (Gulliksen, 1967; Gaudry and Champion, 1962), and three (Stake, 1961) parameters free to describe the individual, autocatalytic function (Woodrow, 1938), negative exponential (Hull, 1943), sigmoid (Culler and Girden, 1951), etc. Although all these curves have some shared similarities, they suggest that probably there is no such thing as a unique function best describing every learning process (Bahrick et al, 1957).

(Tucker, 1959, made an interesting attempt to describe classes of individual learning curves.)

The learning tasks involved in these examinations were usually simple mental tests (such as memorizing names, additions, etc.) or psychomotor tests (like two-hand coordination, discrimination, reaction time, etc.).

The findings of these studies have been quite controversial. Intelligence has been defined as the ability to learn, yet most of the correlations between measures of intelligence and changes resulting from training are reported to be essentially zero. It is widely believed that new knowledge is generally built upon old, and yet many a study has come to the conclusion that those who know the most at the beginning of training profit least from instruction (DuBois, 1965). One used to consider the ability to learn as a one-dimensional trait, and yet gain in proficiency in one task has often been reported to be unrelated to gains in other tasks (e.g. Hall, 1936; Woodrow, 1938; Green, 1953). It would seem that, rather than a unitary learning ability, there exist several relatively independent learning abilities depending upon the kind of task. However, there remains an unsolved question, whether the low correlations between measurements on different tasks are not simply a side effect of the methods (e.g. the scaling used to measure the phenomenon.

18.1.3 Studies with complex learning tasks in the U.S.A.

In the last fifteen years more studies* have been conducted in the U.S.A. with complex cognitive learning tasks. Gagné (1962; 1963; Gagné and Paradise, 1961;

---

* This is partly based on some earlier theoretical papers dealing with learning and individual differences (e.g. Ferguson 1954, 1956; Guilford 1961; Humphreys 1960).

Gagné et al, 1962) and his co-workers, who carried out some very interesting work to see how programmed learning materials can be used in proficiency in mathematics, found a moderately low degree of correlation between rate of progression through a learning programme and measures of general aptitude like intelligence. It is perhaps more interesting that they found this method an especially useful research tool.

Stake (1961), investigating the learning of 12 tasks varying along three dimensions (verbal-non-verbal, rote-relational, and motivational), did not find a general learning factor. Instead he found evidence that learning is related to ability and achievement measures.

Allison (1960), looking for definitions of three learning processes (rote, conceptional and motor learning), factor analysed 13 learning tasks and several ability measures and found that there are learning factors related to abilities and there are others independent of them.

Games (1962), investigating a number of verbal tasks, concluded that learning performance is related to memory. Duncanson (1964), who in his Ph.D. dissertation, administered nine learning tests (concept formation, paired associates and rote memory; each type including one task with verbal, one with non-verbal, and one with figural material) with a battery of ability tests on sixth-grade children found that:

a) measured abilities are related to measured learning performance,

b) learning performance depends upon factors which are independent of ability measures,

c) learning in one task is related to learning in others, and

d) performance in concept formation tasks is not related to the ability measures used or to performances in rote-memory tasks.

Recently Guilford and his co-workers have made some very interesting attempts to study the factorial structure of learning tasks, correlating some of the learning abilities with the abilities involved in Guilford's structure-of-intellect (SI) model (1967a). For example, studying the role of transformation abilities in school learning, they carried out a factor analysis involving 10 transformation abilities and using as criterion measure learning tests called "Studying and Remembering", which consisted mainly of three essays of about 400 words, each followed by a 10-item, 4-choice test. The best predictor factor they found was CMU (cognition of semantic units), followed by MMT (memory for semantic transformation) and CMT (cognition of semantic transformation). Altogether, seven SI factors proved to be significant predictors out of the examined ten (Guilford and Hoepfner, 1971).

In another study examining the relation between the SI abilities to learn a certain kind of concept, it was again found that several factors were involved in both kinds of measurement (Dunham et al, 1966; 1968; Guilford and Hoepfner, 1971).

The results of these studies are more uniform than the ones with simple learning tasks; they seem to be more applicable as well. Yeager and Lindvall (1967), for example, report a study investigating different measures of the rate of learning in classroom situations. They point out that in view of the current activity in the development and trial of a variety of school programmes designed to provide for individual differences among students' learning ability such studies will receive increasing attention; and in addition, opportunities for such studies in the classroom will also be greatly increased because of the number of situations where variations in rate of learning.

are permitted.

18.1.4 Studies with complex learning tasks in Germany.

In the scientific analysis of learning one may elect to study very small changes in behaviour such as the acquisition of a single integral act, which, for all intents and purposes, is merely present or absent. For most practical purposes in education interest does not lie in emphasizing the acquisition of single acts but in the development of complex skills possibly composed of many elements and integrated into far more than the simple sum of such elements. Instead of relating change in behaviour to precisely describable external conditions, concern is with relating complex changes in proficiency to broad and general characteristics of the learner and to major events in the learner's environment (DuBois, 1965).

Some recent studies characterized by this practical emphasis, especially those carried out by Guthke (1969a, 1969b) and his co-workers at the University of Leipzig, show some evidence that complex learning tests can serve in educational studies as useful measurements of learning ability without any factor analysis being carried out.

Guthke (1970) distinguished three kinds of learning tests:

a) longitudinal,

b) practice, and

c) programmed instruction.

He began research in each of these three directions. The longitudinal studies are still in progress. In these he has tried to measure the development of 1000 12-16 year-old students by administering at regular intervals the Raven (1958) Progressive Matrices and the Amthauer (1953) Structure of Intelligence test.

Experiments with the practice method were conducted with pre- and post-tests constructed of the sort of tasks used in traditional intelligence tests. The post-test was found to correlate several times better with school achievement (which was used as a criterion) than the pre-test. This result was especially noticeable after a long delay. For example, the correlation between pre-test, on the one hand, and school achievement half a year later, on the other hand, was 0.59; this correlation dropped to 0.47 when considering school achievement two years later. However, the correlation between post-test and school achievement remained at 0.78 whether achievement was measured half a year or two years later. This could lead us to the hypothesis that a post-test might be a better predictor than a pre-test, and that it is the proximal development indicated in the post-test that is the prediction component.

With the programmed learning test method two studies have so far been carried out. In the first study grade one students had to learn some regularity in a classification task; the experimenter told the subjects whether their attempt was right or wrong. The results in this learning test had a 0.65 correlation with mathematics achievement.

In the other, more sophisticated study, the Raven Progressive Matrices were presented by a teaching machine. As in the original test the subjects could choose from given alternatives but in the Guthke study after each decision the machine told the subject whether he had chosen the correct answer and, if not, what sort of mistake he had made. The results with grade six pupils were quite encouraging. While the correlation between the pre-test (the original Raven test) and school achievement was only 0.39, the sum of the pre- and post-test results had a much stronger correlation (0.62) with school achievement. Calculating a gain score, using equation 1 they got a surprisingly high correlation (0.82) between the gain scores and the teachers' intelligence rank order. (For further details see Guthke, 1972).

$$S = \frac{X_2 - X_1}{X_{max} - X_1} \qquad \text{Eq. 18-2}$$

where S = gain score

X_1 = score in pre-test

X_2 = score in post-test

X_{max} = the maximum possible score in the test.

In a series of school tests, Schröder (1968) developed a combined intelligence and learning test (KLI 4^+). The first part of this test has the same title ("Geheimschrift") as our test. The main differences are that while Schröder's test consists of a relatively short learning material (1 page) and several tasks (similar to traditional intelligence tests) our test has a longer learning material and only 1 task measuring the effectiveness of the learning process; secondly that we used a linear programme in the learning material; and thirdly that we included in our measurement the creative application of the acquired information.

18.1.5 Intelligence and learning ability

One major question arises in these studies again and again: What is the relationship between intelligence and learning ability (Glaser, 1970). We once more emphasize that for many psychologists, involved in the development of intelligence tests, a strong relationship between these two variables seems to be evident. However,

> Theories of intelligence are often vague or inconsistent about whether MA, IQ, or both, should be related to learning. According to one view, it is MA score that is related to learning, because MA is a measure of the developmental level of the organism, and it is this level which limits learning ability. In these terms, IQ is not a measure of present learning ability, but rather a mathematical statement of the rate at which learning ability will change in time. An alternative view classifies the intelligence test as an achievement test of universally taught subject material. The MA score is then a measure of how much has been learned. If training is assumed to go on uniformly in time, the ratio MA/CA is the slope of a life-time learning curve, characteristic of the organism and recoverable in new, miniature learning situations. Consequently, IQ is the measure of present learning ability (House and Zeaman, 1960).

Perhaps Woodrow's (1933c) findings first questioned whether ability to learn, based on laboratory and classroom experiments, in the sense of ability to improve with practice, is identical with intelligence. Since then, several correlational studies (e.g. Stake, 1961) seemed to support Woodrow's opinion and several others seem to contradict it (e.g. Duncanson, 1964). For a better description of problems and literature on the relation of learning with special regard to retardation, see Zeaman and House (1967), Denny (1963) and Lipman (1963).

Recently, Jensen and his co-workers (Jensen, 1961, 1963, 1968a, b; 1969a,b; 1970; Jensen and Rohwer, 1968; Rapier, 1968) argue that although there is a relationship between learning ability and intelligence, it is a function of the subject's socio-economic status (SES; see Fig. 18-1).

Without having really strong evidence, while constructing this test, we believed that the failure to find a correlation between the learning ability measurement and the traditional intelligence tests probably was a by-product of the method the researchers used (see e.g. Ruch, 1936; 1961). Our data (see next section) seemed to support this belief.

18.1.6 Personality variables and learning ability

Very little literature exists on the relationship between "motivational" or "personality" individual difference type variables and learning ability. The work of Eysenck (1965) and Taylor (1958) are some early examples, although confined to limited learning phenomena and a limited range of "personality" variables. The proliferation of studies using the "Taylor anxiety" scale illustrates the relative lack of sophistication of learning psychologists in this area. Witkin et al., (1954) have been attempting to explore some "cognitive style" variables as possible "personality" type variables interacting with performance at different stages of learning, and recently, Jenkins (1967)

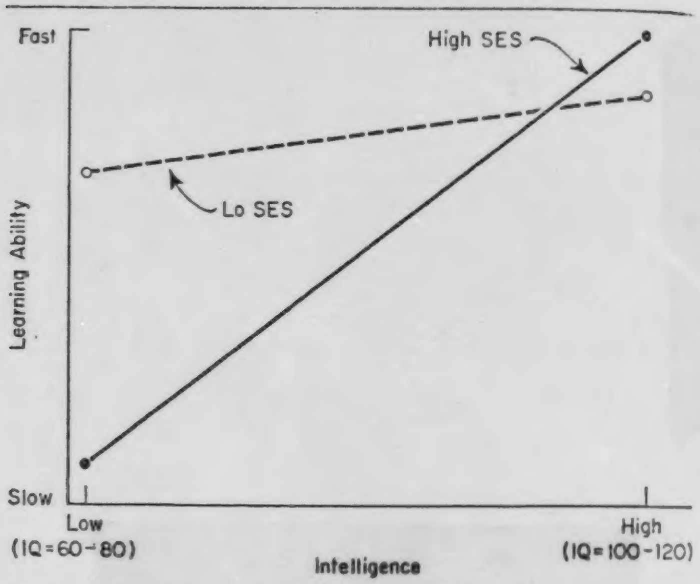


Fig. 18-1. Summary graph of a number of studies showing the relationship between learning ability and IQ as a function of socio-economic status. Jensen, 1970.

emphasized the need to look at motivational or "personality dynamic" types of variables in relation to learning. However, personality tests of "rigidity-flexibility", "anxiety", "extraversion", etc. would fail to predict individual differences in associative interference, level of performance, or rate of learning at any stage of practice (see e.g. Fleishman, 1967).

18.1.7 Our own studies with programmed learning material in Hungary

A MODEL FOR TESTING LEARNING ABILITY

Our own interest in measuring learning ability began with an awareness of the growing need to improve the college admission system in Hungary. Numerous investigations have shown that neither matriculation, nor the present entrance test measuring the students knowledge, is satisfactory for this purpose. The problem is not limited to Hungary and, throughout the world, it is difficult to envisage the next step in its solution (Whitla, 1969). As a part of the admission procedure in Hungary (Biczók et al., 1967) as well as in other parts of the world (see, for example, Alexakos, 1966) there have been some not too successful attempts to include a kind of traditional intelligence test. To try to find a more suitable measurement for this special purpose we turned to the Vygotsky (1962) theory on "the next zone of development" which emphasizes the results of learning experiments for both diagnostic and prognostic purposes (discussed above).

The following model (Klein, 1970a) for measuring ability to acquire new information and use it creatively was developed:

1. The first step of the testing procedure should be the measuring of the knowledge level of the investigated individual, using standardized achievement tests. With the help of these the

$$t = \frac{h}{H} \qquad \text{Eq. 18-3}$$

index shows the subjects' knowledge level, where

h = the quantity of information that the investigated individual has on the material in question, and

H = the quantity of information defined as belonging to the material.

2. After grouping the subjects according to their knowledge level, each group must be made homogeneous in terms of knowledge level by teaching it to a definite level.

3. An instructional programme (the learning task) must be assembled for each group. The use of a computer-run teaching machine with a flexible programme was suggested.

The first factor which influences the measures of information acquisition and creative application (IA for future references) is the time (T) needed to acquire information in the learning task. It is our opinion that the time required for learning influences, though perhaps not decisively, the level of intelligence. Although it would be difficult to prove this statement in terms of intelligence in general, in the case of our entrance examinations such a position is obviously justified, in view of the increasing quantity of knowledge a student is expected to acquire in shorter and shorter time periods.

4. At the end of the studying process (after the last step of the programme), the knowledge acquired is measured by giving the subject such tasks as can be solved on the basis of the information contained in the programme. The proportion of the quantity of information produced in the tasks solved (h_f), compared to the quantity of information required by the task (H_f), we consider as the measure of information acquisition:

$$v = \frac{h_f}{H_f} \qquad \text{Eq. 18-4}$$

5. It was felt that as a major disadvantage of the traditional intelligence tests is their exclusion of creativity, a measurement of the ability to creatively use the information gained should be included in our test. Torrance (1962a,p. 132) reports generally very low (between -0.02 and 0.32) correlation coefficients between creativity and intelligence measurements. To exemplify that investigators have been insisting for years, to little or no avail, that tests of creative thinking should be used together with a reproductive test such as traditional measures of intelligence to identify and guide the development especially of intellectually talented individuals, he cited Simpson, who presented this viewpoint as early as 1922.

In 1970 Getzels made a strong case in connection with selecting students for college entrance: that is, we have to give a chance not only to those students with "convergent" intellectual ability but also to superior "divergent" students.

For this purpose we built into our model the construction of new problems, concepts, or systems inspired by the programme but not found in it. The method of measuring creativity (k) would of course depend on the type of programme used. In some cases k could be considered the proportion of the new information produced (H_a) to the information provided by the programme (H_p):

$$k = \frac{H_a}{H_p} \qquad \text{Eq. 18-5}$$

6. From T, V and k we have to construct the index of information acquisition and creative application (IA), by assigning a suitable weight to each of these three variables. The IA index should be a sum of two separate items, the measurement of information acquisition and the measurement of creative application. The measurement of information acquisition is obvious a direct function of V and inversely proportional to T.

From these theoretical considerations we proposed as a measurement of IA the following equation:

$$IA = \frac{V^x}{T} + y\,k \qquad \text{Eq. 18-6}$$

where x and y > 0 are weights to be defined by maximizing the validity of the test (see also sec, 18.2.1 and 18.3.1).

For practical reasons we have not until now carried out any studies in college admission based on this model. Studies at a lower level, based on simplified versions of parts of this model are summarized in this and the following section.

INTELLIGENCE, SUBJECT KNOWLEDGE, DILIGENCE

In an earlier study (Klein, 1970a, 1973e) we began from the assumption that the knowledge level a student attains in a school subject will depend on his diligence on the one hand and on his intelligence or learning ability on the other. It would seem, however, that the connection between these three factors should be different for different subjects; that is, success in some subjects should depend on diligence more than intelligence (learning ability) while in other subjects it should be the opposite. We were looking for a test which could make this discrimination; that is, give different correlations with knowledge level and intelligence in different subjects.

Dividing them into ten groups, we examined 150 pupils (aged about 16) coming from a school for training skilled workers. The first chapter of the Hungarian translation (1968) of Carmen's *Programmed Vector Algebra* served as learning material. By giving an initial test we got fair evidence that we could consider the knowledge level of the subjects in this topic as zero though they possessed all the knowledge necessary to follow the programme. Therefore, as an indication of learning ability, we considered only the results of the post-test. On the basis of the model described before the following formula was invented to score the learning test:

$$IA = \frac{N^2}{T} - \frac{m}{10} \qquad \text{Eq. 18-7}$$

where IA (information acquisition and creative application)

IA = the score for the learning test

N = the number of tasks concerning the learning material resolved in the post-test

T = the time in minutes to acquire the given information

m = the number of tasks constructed by the subject, based on the information supplied by the programme but not included in the programme or the post-test.

The empirical weights in this formula seemed to have the advantage of roughly balancing information acquisition $\left(\frac{N^2}{T}\right)$ and creative application $\left(\frac{m}{10}\right)$.

The IA scores varied between 0 and 1.8 (with a mean of 0.55); the value of N was between 0 and 4 (with a mean of 2.8); that of T was from 10 to 89 minutes (with a mean of 36 minutes) and m varied between 0 and 5 (with a mean of 2.5).

To study the correlation between this learning ability test and other measures of student intelligence we gave the Hungarian form of the Wechsler test (MAWI; Kun-Szegedy, 1971), and also asked teachers to rank the pupils according to their intelligence. The use of Kendall's concordance coefficients showed that, out of 10 teacher groups participating in this investigation, 7 groups ranked their pupils congruently. (Each teacher group consisted of 4 teachers in the subjects of practical education, technical instruction, mathematics and specialized drawing. The concordance coefficients ranged between 0.26 and 0.82, the lowest significant coefficient being 0.45.) We also asked the teachers to rank the students according to their diligence and their subject knowledge. Some of the results were the following: The rank correlation coefficients between the teachers' estimations of student diligence and knowledge of the subject matter were around 0.7. The teachers' rank orders of intelligence and knowledge of subject matter correlated at around 0.7 as well, and intelligence and diligence rankings correlated at around 0.6. There were no significant differences between the correlation coefficients in different subject matters.

The IA scores correlated at around 0.4 ($p < 0.05$) with the teachers' intelligence rank orders as well as with the Wechsler test results. They had a very low correlation with the teachers' ranking of student diligence (around 0.15), and a low correlation with the teachers' ranking of subject matter knowledge (around 0.23). But the IA correlations were significantly different according to the different subjects, which was not the case with the MAWI test or with the teachers' rank orders. Concerning only the connection between IA and subject knowledge, the coefficients were 0.36 and 0.31 for mechanical drawing and mathematics (which subjects are probably closely connected to intelligence) and 0.02 for technical training, which has probably not too much connection with it. There was a successful discrimination, therefore, between subject matters when correlating knowledge level with learning ability as measured by the IA index - despite the fact that without a factor analysis, we cannot say what learning ability factors were present and with what weights.

THE USE OF THE TEACHING MACHINE IN THE MEASUREMENT OF LEARNING ABILITY

Some years ago, Sorenson (1963) suggested that the concept of intelligence be abandoned in favor of a measure of learning ability derived from performance in a teaching machine learning situation. But as Duncanson (1964) says: "Until such a measure can be shown to relate to some meaningful criterion and to have some degree of generality, however, intelligence will remain the more useful concept."

In the second study (Klein, 1972b), we made an attempt to relate a measurement of learning ability derived from performance in a teaching machine learning situation to measurements of intelligence and creativity.

As a part of the psychological examination of 3 classes taught with the OPI Mathematics Project's method for as long as eight years, we carried out a learning test again using the first part of Carmen's Programmed Vector Algebra presented by a Grundy Tutor teaching machine. In this case we have a pre-test concerning the knowledge necessary to study the programme. It consisted of 48 scorable parts.

A parallel form of this test served as post-test. The Kuder-Richardson "Formula 21" (see, for example, Magnusson, 1966, p. 116) estimated reliability coefficients of 0.83 for the pre-test and 0.78 for the post-test. According to Equation 18-1 the 0.39 correlation between the pre-test and the post-test would result in a 0.3 reliability coefficient if a simple gain score were used. So it was again decided to use more sophisticated methods.

As in the preceding study, there was an additional task that of constructing new questions based on the studied material but not existing in the post-test. We considered this task as a kind of creative application of the existing material. For this test, however, we grouped the students' questions according to general type and considered the inverse of the number of questions in each group in order to assign originality scores; a subject's originality score was the sum of the scores of his answers; the number of questions invented by one subject was his fluency score; and a kind of combination of these two was considered a parameter of creative application.

To consider the learning process we used a technique borrowed from information theory. V_p (the "programme utilization ratio" of a student) was defined as the following:

$$V_p = \frac{h_p}{H_p} \qquad \text{Eq. 18-8}$$

where h_p = programme utilization (i.e., the amount of the programme which a student found it necessary to follow in order to attain the required knowledge level), and

H_p = the maximum possible programme utilization.

Furthermore,

$$h_p = \Sigma\, p_i h_i \qquad \text{Eq. 18-9}$$

where p_i = the probability of the students giving this particular answer to question i.

This probability is estimated by

$$p_i = \frac{f_i}{n} \qquad \text{Eq. 18-10}$$

where f_i = the number of subjects giving this particular answer to question i

and n = the number of subjects.

h_i = the amount of information gained by the student from question i.

This can be defined, according to the basic function of information theory, as

$$h_i = \log_2 \left(\frac{1}{p_i}\right) \qquad \text{Eq. 18-11}$$

As a measurement of the whole process of information acquisition and creative application by teaching machine the following index was used:

$$IA_{TM} = \frac{(5U)^2}{E.V_p.T} + 5m \qquad \text{Eq. 18-12}$$

where IA_{TM} = information acquisition and creative application measured by teaching machine

U = score in post-test

E = score in pre-test

T = the time of the information acquisition process

V_p = the programme utilization ratio explained above

and m = measurement of creative application, also explained above.

In spite of the more sophisticated information theory method used in this study it is easily seen that equation 11 is still based on the model described before and is derived from Eq. 18-6 by still using arbitrary weights.

In an attempt to validate the IA_{TM} measurement we used more traditional tests, hoping to find a reasonably high correlation of our measurement of learning ability with intelligence and creativity. We used two intelligence tests (Raven, 1958 and MAWI, the Hungarian version of the Wechsler test) and four creativity tests (Consequences (remote), Unusual (alternate) uses, Number association (uncommon), Make-a-figure Test (shifts); see Guilford-Hoepfner, 1971). For the creativity tests we used an unusual scoring method (see Barkóczy-Klein, 1968).

Some of the results were the following:

the correlation between IA_{TM} and the Raven test was 0.55;

the correlation between IA_{TM} and the MAWI test was 0.59;

the correlation between IA_{TM} and the combined score of Raven and MAWI was 0.57.

The number of subjects was 23. All of the correlations were significant at the 5% significance level. Two of the creativity tests correlated significantly at the 5% level with IA_{TM} scores: Consequences (remote) (0.69) and Unusual (alternate) uses (0.44). Two others did not correlate significantly: Number association (uncommon) (0.09) and the Make a figure test (shifts) (0.17). These correlation coefficients were based only on the results of 19 students involved in both the learning ability and the creativity measurements. The results fairly coincided with our expectations of validity.

Concerning the recent evaluation of the OPI Mathematics Project in terms of learning ability, we were permitted to use only the eight most intelligent members of a control class who were matched with experimental students in terms of the average of the Wechsler and Raven I.Q's. Although the mean of the IA_{TM} scores of the experimental group was better than that of the control group, the difference was not significant (using the Wilcoxon test). presumably because of the small number of students. Nevertheless, we feel that further use of this method with larger groups could prove fruitful.

Both studies gave some encouraging results in spite of the lack of specially designed programmed learning material. However, we thought that to go further in this direction we would certainly have to develop several programmed learning materials constructed to involve all the different kinds of thinking (abstraction, generalization, etc.). The Secret Writing Test is a first step in this direction.*

18.2 The Budapest Study

18.2.1 Method

Substantially the same method was used in the Budapest, Sherbrooke and New York Studies. We therefore describe the method in detail only here and shall confine our account to relevant differences in the sections dealing with the other two Studies.

The Test is titled

"Secret Writing, A Test of Information Acquisition and Creative Application"**, and consists of two parts.

Part 1 is programmed. While the child goes through pages 2 - 43 in the test booklet he learns a secret code for converting some vowels into some special signs and back in context. Fig. 18-2 shows an example out of the 21 exercises. Page 44 consists of a riddle (Fig. 18-3). The children have to "translate" it into the secret code just learned. Its purpose is to measure the amount of information the children have acquired.

---

* Recently we began to develop a new series of learning ability tests based on group structures (Dienes & Klein, 1973).

** see Klein 1969 for the full title of the Hungarian original.

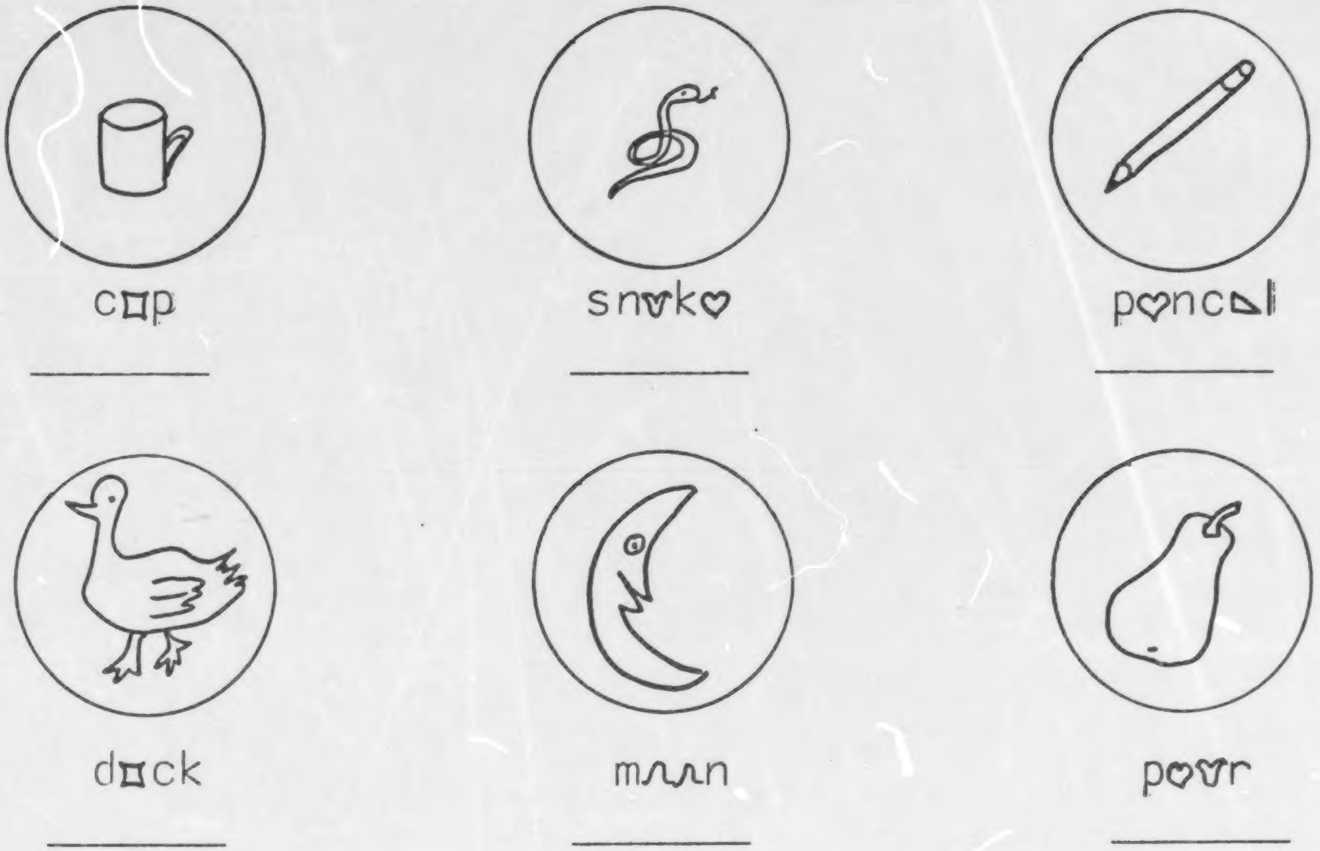
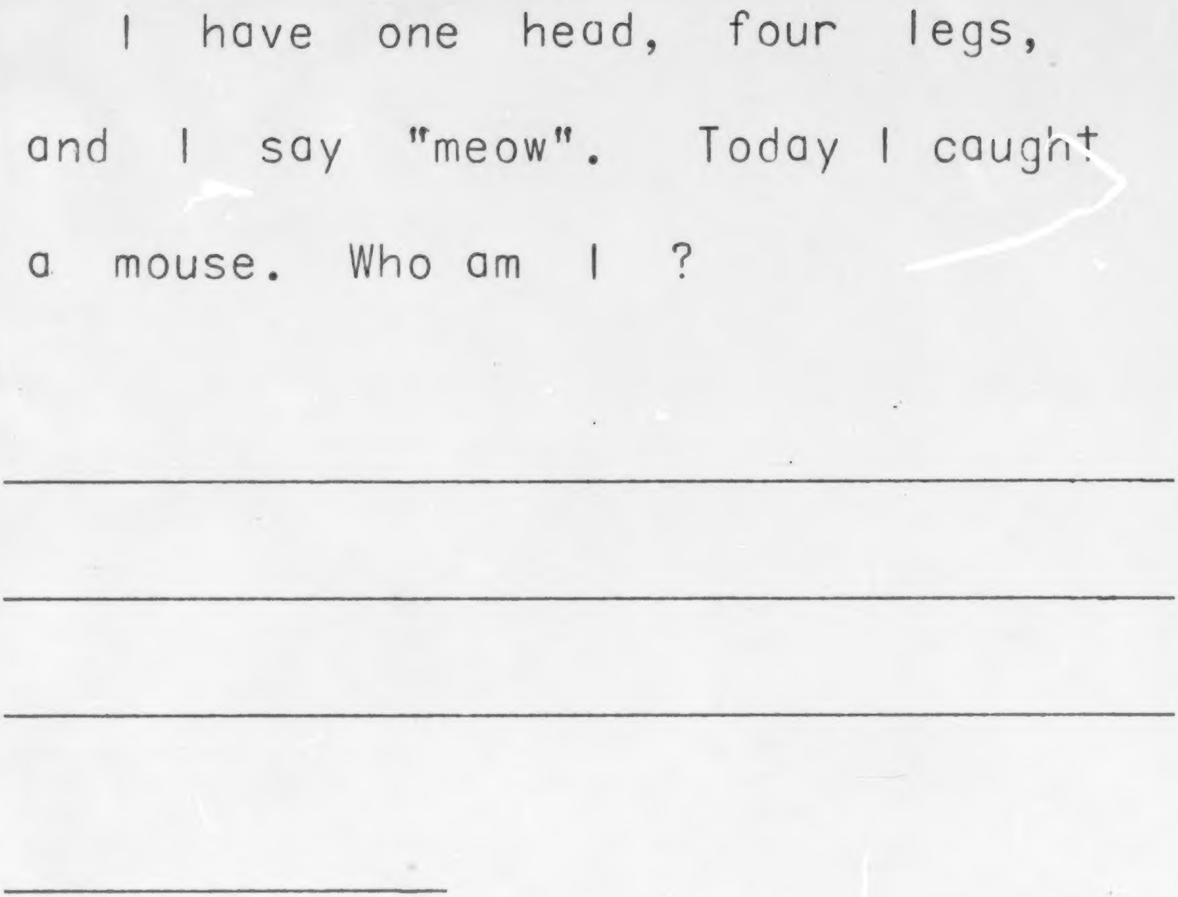

c□p
sn▽k♡
p♡nc▷l
d□ck
mᴧᴧn
p♡▽r

Fig. 18-2

"Guess what is written."

A translation exercise from the test booklet.



I have one head, four legs, and I say "meow". Today I caught a mouse. Who am I ?

Fig. 18-3

"Translate and answer this riddle in secret writing."

The last translation exercise from the test booklet

Because the Secret Writing Test requires written responses, administration to groups of pupils is possible, although individual administration is recommended to younger and/or disabled children.

The actual pupil working time until the end of the first part of the test (the learning time) varies between 15 and 80 minutes. Most students finish in about 40 minutes. Students usually spend 5-10 minutes on the second part of the test (creating new secret codes). To cover preliminary instructions, handling out of booklets, etc., no less than 70 minutes free of interruptions should be scheduled for the test. It should, however, be possible for any student to continue working longer without feeling any pressure.

ADMINISTRATION OF THE TEST

The following directions are to be used when presenting the learning ability test to a group of children:

1. PRELIMINARY REMARKS. Begin with whatever introduction is appropriate to the group and testing situation; this might typically include an explanation that the test is "to show us how children of your age learn."

2. DISTRIBUTE THE TEST BOOKLETS.

3. PERSONAL INFORMATION. Have the children fill in the information asked for on the front cover, then say:

4. INSTRUCTIONS

"Good. Put down your pencils and I'll explain what you have to do.

"Today you are going to learn a new way to write -- how to produce some secret writing. In the booklet there are several exercises which will teach you secret writing. Do the exercises, and at the end there will be a riddle for you to change into secret writing.

"As you go through the booklet, you will find an exercise written on one page -- and then you will find the answers to the exercise on the next page. Try to answer each question by yourself and then, when you have found an answer, compare it with the correct answer which is given on the next page. If you find that you have made a mistake, try to understand by yourself, and correct your answer if you like. That way you'll learn secret writing well and you will be able to answer the last question even though no correct answer is given for it.

"You'll have to learn secret writing all alone -- I cannot help you -- but you'll see, it's fun. Do all the exercises from the beginning of the book until the end, and work as fast and carefully as possible. You may work for as long as you need to. When you finish the last question, raise your hand and I'll go to your desk."

"Are there any questions now? (ANSWER ANY QUESTIONS.)

"Good. Turn the page and start."

5. BEGIN TIMING.

6. AFTER A FEW MINUTES SAY:

   "Don't forget that the correct answer is always written on the page after the question."

7. MARK THE TIME. As each student finishes the last question of Part I, arrives at page 45, and raises his hand, mark in his booklet the time it has taken him. Then turn to page 46 in his booklet and say: "Good. Could you read this and then do what it tells you to do?"

8. GO BACK TO EXPLAIN INDIVIDUALLY WHERE NECESSARY.

9. FINISH. Give each student as much time as he needs (to a maximum of $1\frac{1}{2}$ hours for the entire test) and then collect his test booklet.

SCORING

The learning ability test gives three partial scores and a combined score for each subject; the partial scores are the following:

(i) Learning Time (T)

The learning time is the time in minutes (T) a student uses to finish the first 45 pages. This usually varies between 20 and 80 minutes.

(ii) Learning Score (N)

The learning score (N) is calculated from the student's translation of the riddle and its answer. Each vowel is considered separately; if it is correctly translated at least once, $\frac{1}{2}$ mark is given; if it is correctly translated each time it is used, a total of 1 mark is given. Summing the scores for the six vowels, the range of possible values of N is 0 to 6.

This method is used even if the subject has misunderstood the question and has, for example, translated only the riddle or only the answer:

e.g. "the cat" gives him a score of 2.

No attention should be paid to the quality of the child's drawing.

(iii) Creative Application Score (m)

The score for creative application (m) is the sum of the scores which are given to each different secret writing invented by the subject (m_i).

$$m = \Sigma m_i \qquad \text{Eq. 18-13}$$

Omitted from the scoring are those secret writings which do not show a minimum of:

a) secretness, e.g. transformation into printed letters only, in cursive or printed styles.

b) invention, e.g. those which are only a repetition of the secret writings learned before.

c) difference, i.e. are a repetition of secret writing, for which the subject has already received a score.

Every secret writing which fulfills minimal requirements of secretness, invention and difference will receive a creative application score m_i, which consists of two partial scores: the elaboration score m_e and the originality score m_o.

$$m_i = m_e + m_o \qquad \text{Eq. 18-14}$$

(iv) Elaboration Score (m_e)

This score refers to how well elaborated or how exhaustive a secret writing is:

m_e = 0.5 for minimal attempt to change the secret writing as learnt.

m_e = 1.0 for new secret writing worked out partially

m_e = 1.5 for new secret writing worked out completely.

(v) Originality Score (m_o)

A secret writing will be considered as usual and will not be rewarded for originality if it is essentially a one-for-one replacement of letters by new symbols taken letters from a list of mathematical symbols such as

($>$, $<$, +, -, $\div$, 1, 2, 3...) or (○, △, □).

In any other cases the secret writing will be rewarded with one point for originality.

Of the numerous possible combinations of T, N, and m to give a combined score of learning ability and creative application (IA), in the Budapest study we used the following function based on Equation 18-6 with empirical weights :

$$IA = \left(\frac{N^2}{3T} + \frac{m}{9} \right) \times 100 \qquad \text{Eq. 18-15}$$

18.2.2 Treatment of data

The differences between the IA scores of experimental and control classes were examined using the Mann-Whitney test. The Spearman rank correlation coefficient was used to compare the IA results with the Binet test.

18.2.3 Expectation

Higher learning scores in experimental classes than in control classes would show greater development of the students' mathematical learning ability, thus supporting hypothesis 13.

18.2.4 Results

IA indices were higher in the experimental classes in both grades 4 and 5 (47.8 and 38.3 in grade 4 and 55.1 and 37.8 in grade 5), the difference between the experimental and control class was significant only in grade 5.

18.3 The Sherbrooke Study

The Sherbrooke Study differed from the Budapest Study in the Method of Scoring, the Treatment of Data, and the Results.*

18.3.1 Scoring

The partial scores were defined substantially in the same way as in the Hungarian study. The combined score of information acquisition and creative application was defined with the following equation:

$$IA_{raw} = \frac{N^3}{T} + n \qquad \text{Eq. 18-16}$$

(The index "raw" remains for differentiation with those IA scores converted into I.Q. equivalents - see later.)

This equation was derived from the model described in 18.1.6 (see Eq. 18-6) but to define the weights of the variables a much more sophisticated method than in the Budapest study was used. It involved an optimalization process using the following procedure:

---

* see Klein 1972c for the full title of the French version of the Secret Writing Test.

Besides the Lorge-Thorndike I.Q.s measured when the students began grade 1, the results of three other intelligence tests were available as the Raven, Otis-Lennon, and Cattell Culture air tests were administered to the same population at practically the same time as the Secret Writing test. Under these conditions we supposed that the best criterion for optimalizing the weights of the three variables constructing the IA would be the average of the results of these three intelligence tests; the three tests therefore had to be scored on the same scale, for example, an I.Q. - equivalent scale like the Wechsler Intelligence Scale (WISC; Wechsler, 1949) which has a normal distribution, a mean of 100 and a standard deviation of 15. Since French-Canadian norms were not available for any of the intelligence tests it was thought advisable, in order to get as valid a combined intelligence score as possible, to do our own norms for two of the tests (Raven and Cattell) and to refer for the third (Otis-Lennon) to American norms which have the advantage of being based on a much larger sample and therefore give better differentiation in age groups (3-months difference). The normalization of the Raven and Cattell tests was done in grades 4 and 5 separately. (For a discussion of normalization see Magnusson, 1966, pp. 235-238.) The normalized scores were transformed linearly into I.Q. equivalents and the average of the three I.Q.s was calculated.

The optimalization procedure itself then consisted of the calculation of the correlation coefficients between the average of the three I.Q.s on the one hand, and the IA's calculated by experimenting with different weights (x and y) from equation 18-17 on the other hand:[1]

$$IA = \frac{N^x}{T} + ym \qquad \text{Eq. 18-17}$$

(Equation 18-17 is a simplification of equation 18-6 replacing the information theory measurements V and k with the less sensitive N and m which are found simply by counting responses as described earlier.)

---

[1] The computations were carried out in the computer centre of the Université de Sherbrooke by Jean-Marie St. Pierre.

Using this procedure to find an optimal combined score we indeed emphasized the relation between the IA and the traditional I.Q., though we supposed that in our measurement besides the ability to learn (intelligence) the creative application of the gained information is also included. As Torrance Tests of Creative Thinking (1966, 1972) and the Secret Writing test were carried out on the same populations in the Budapest and Sherbrooke Studies, we could use a combined score of the intelligence tests and the Torrance tests as a criterion variable, but in this case we would have a circulus vitiosus with our weighting problem; we should first have had to find appropriate weights for the components of the criterion variable. To avoid this it was decided to verify the correlation between IA and the Torrance tests after defining the weights of the components of the IA. (See section 18.1.7.)

It should be pointed out that equation 18-17 is statistically quite a poor formula because it simply adds weighted raw scores, without normalizing the variables. Even if we suppose that $\frac{N^x}{T}$ (the scores on the part of the test measuring information acquisition) has a normal distribution, there is no reason to suppose this about the creative application scores. Consequently there is no reason to expect a normal distribution for the IA, which fact could fairly lower the correlation coefficient between IA and I.Q. Because of this a correlation coefficient around 0.5 was expected and would be considered quite satisfactory. The results of the optimalization procedure are shown in Fig. 18-4.

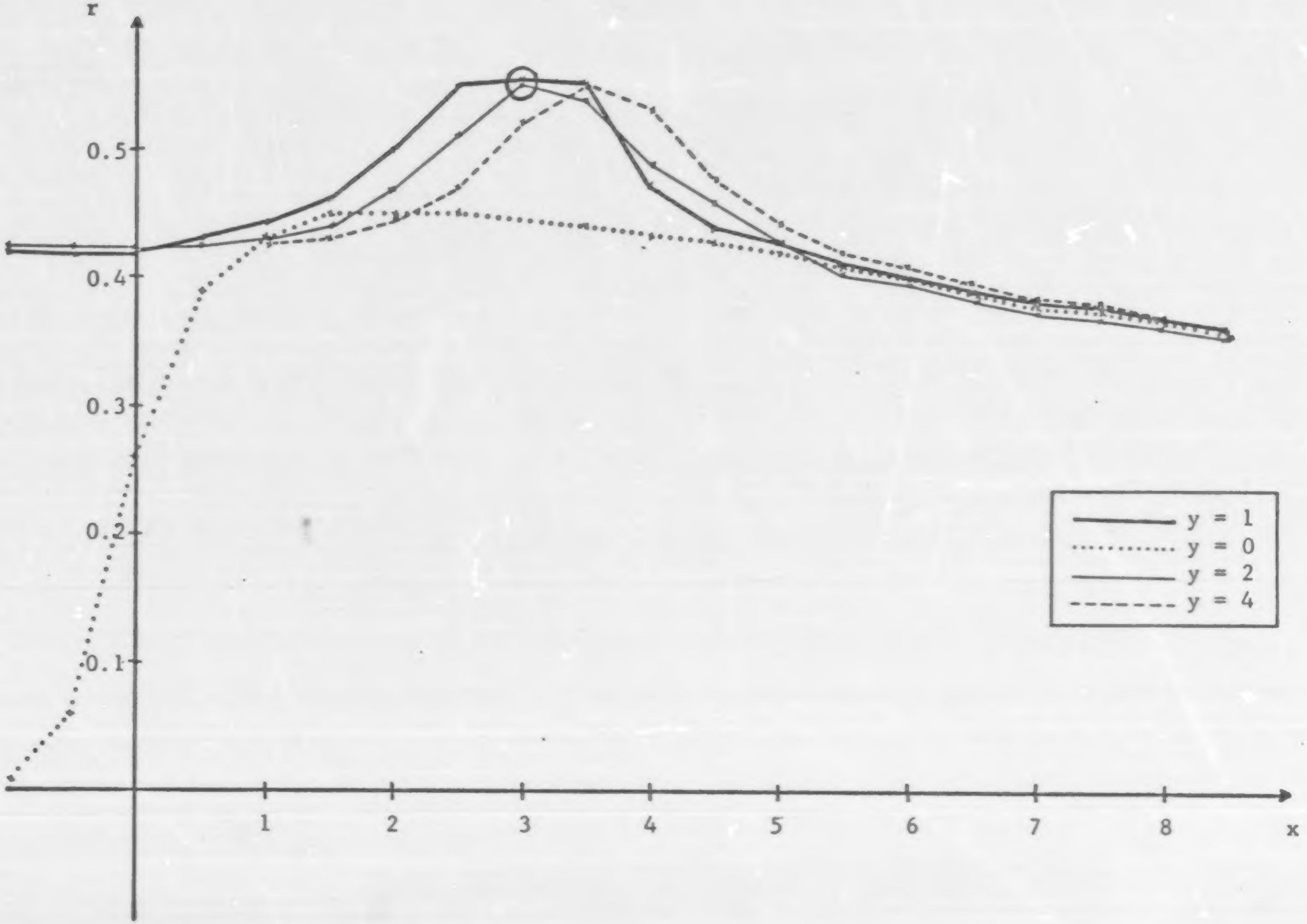


Fig. 18-4 The correlation between IA = $\frac{N^x}{T} + ym$ and IQ (as a mean of three different intelligent test scores) as a function of the weights x and y (a circle shows the optimal combination of x and y.)

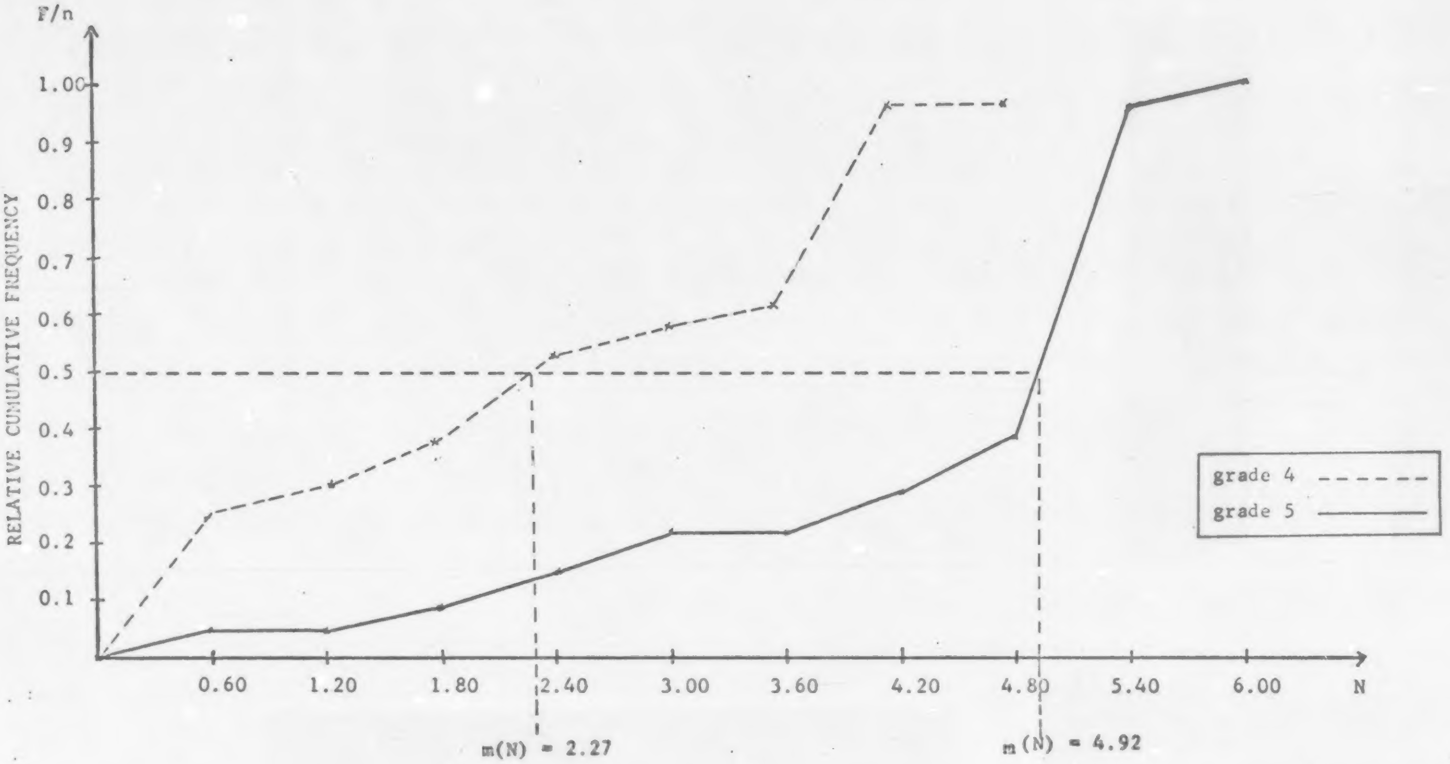


Fig. 18-5. The cumulative frequency distribution of the number of acquired code letters (N) in the Secret Writing Test in grades 4 and 5, with their medians m(N).

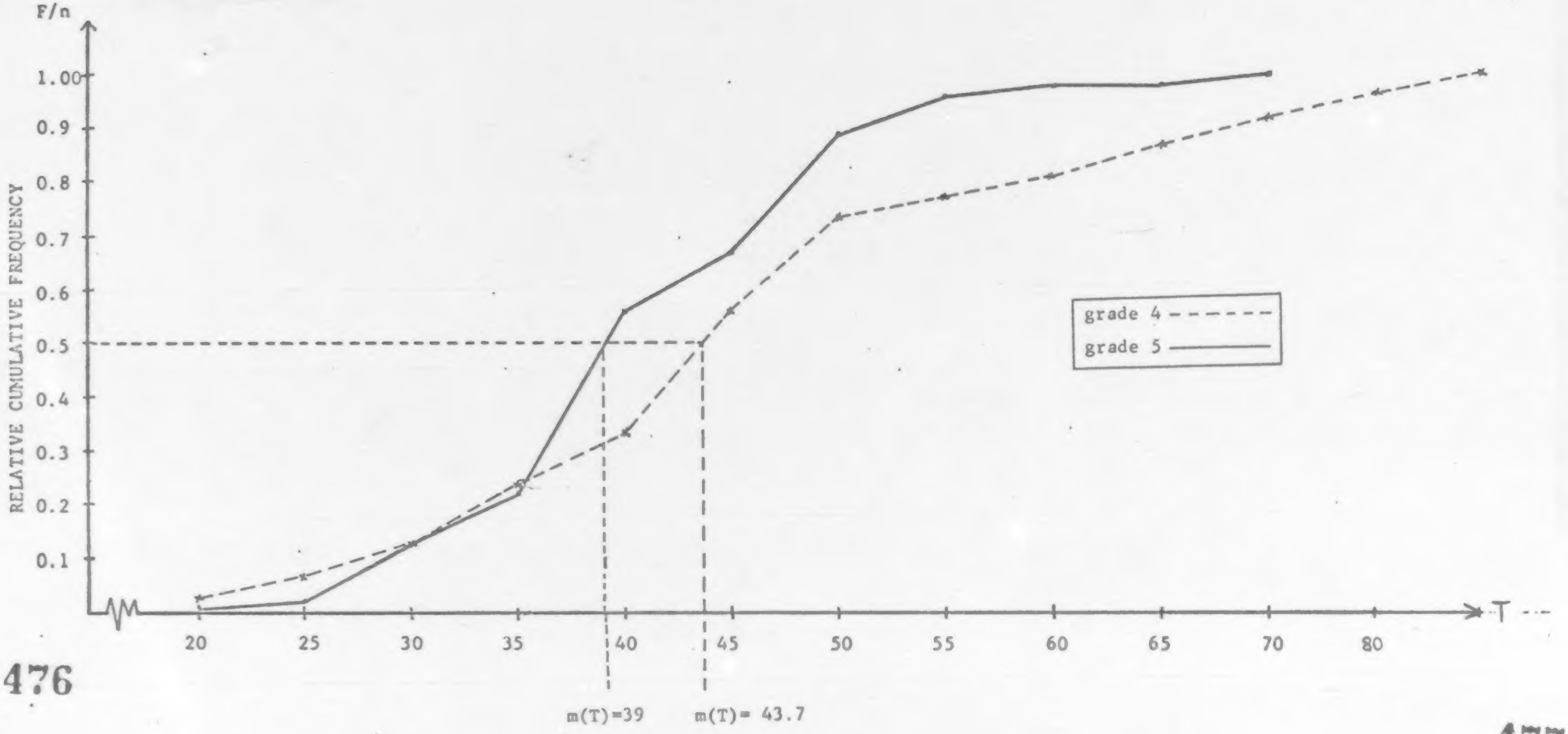


Fig. 18-6. The cumulative frequency distribution of the time (T) spent on the Secret Writing Test in grades 4 and 5, with their medians m(T).

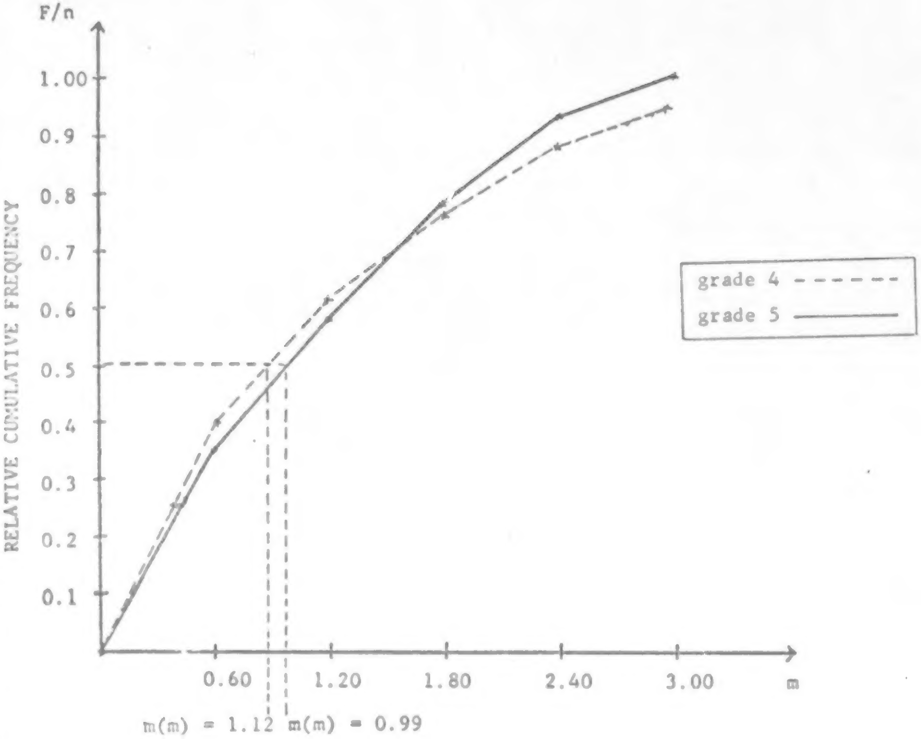


Fig. 18-7 The cumulative frequency distribution of the number of new codes invented by children (m) in the Secret Writing Test in grade 5 and 4 with their medians m(m).

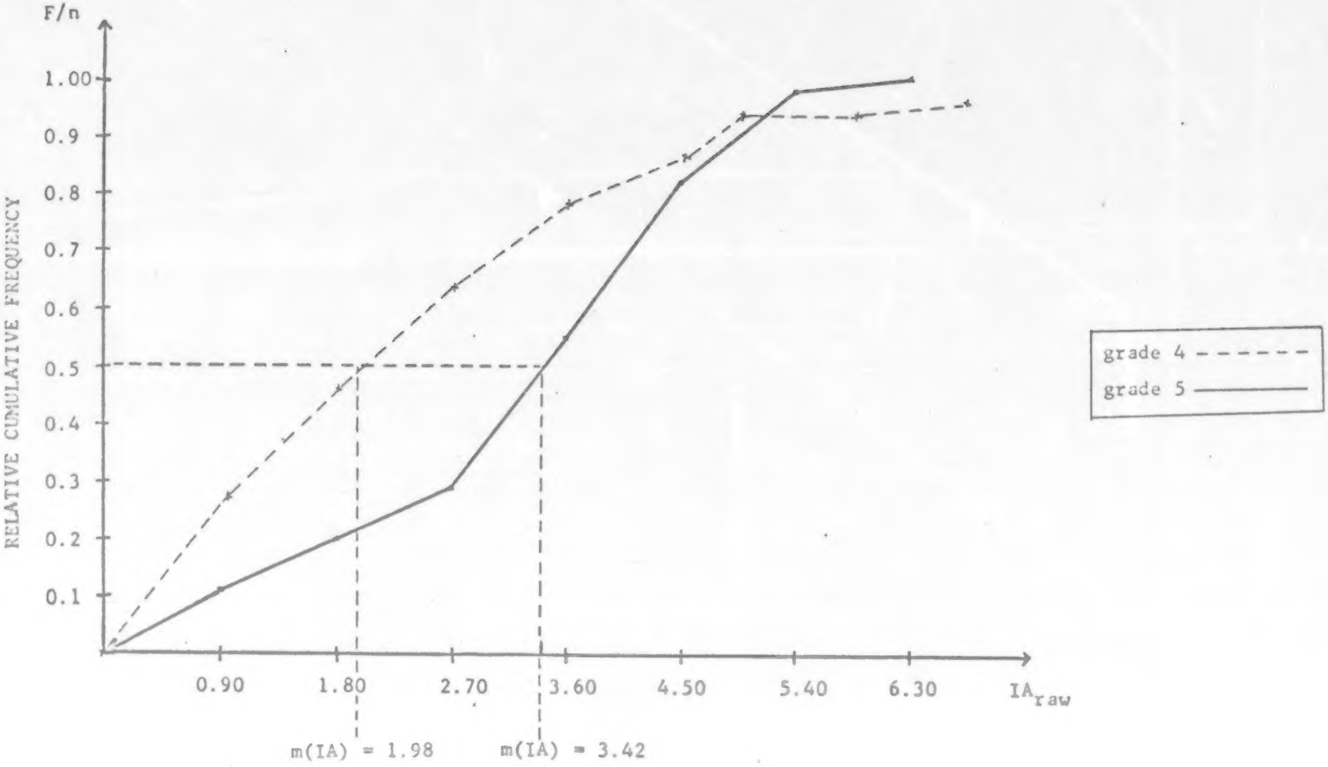


Fig.18-8 The cumulative frequency distribution of the IA scores in grade 5 and 4 with their medians m(IA).

Negative values were also tried for x and y; as expected, these gave poor correlation coefficients. It is interesting to notice that, as we had supposed in the model described in sec. 18.1.7, it was necessary to balance information acquisition $\left(\frac{N^x}{T} \right)$ and creative application (ym). As can be seen from Fig. 18-4 we found the maximal correlation between IA and the criterion variable when x = 3 and y = 1, yielding Equation 18-16 at the beginning of this section.

The cumulative frequency distributions of

N, T, m, and IA in grades 4 and 5 are shown in Figures 18-5, 18-6, 18-7 and 18-8.

To see that both $\frac{N}{T}$ and m have preceptible effect on IA we used a simple statistical procedure:[1]

If m(N), m (T), and m(m) are medians, $m_{25}(N)$, $m_{25}(T)$, and $m_{25}(m)$, the 25th centiles of N, T, and m respectively then the following equation could be calculated:

$$A = \frac{m(N)^3}{m(T)} + m(m) \qquad \text{Eq. 18-18}$$

$$B = \frac{m(N)^3}{m(T)} + m_{25}(m) \qquad \text{Eq. 18-19}$$

$$C = \frac{m_{25}(N)^3}{m_{25}(T)} + m(m) \qquad \text{Eq. 18-20}$$

where A is the so-called pseudo-median of the values IA.

---

[1] The method described here originated in a discussion with B. Connes in February, 1973. We wish to express our thanks to him while nevertheless emphasizing that he should not be held responsible for any interpretation found herein.

As (A - B) is a certain measure of the effect of m on IA and (A - C) is the same kind of measure of the effect of $\frac{N}{T}$ on IA. The ratio

$$D = \frac{A - B}{A - C} \qquad \text{Eq. 18-21}$$

is a measure of how well balanced the effect of these two variables is . (If for example this ratio is too small, e.g. 0.1 or smaller, it would mean that if our weights are $x = 3$, $y = 1$, then m has practically no effect on IA). From practical consideration we expected that D should be

$$\frac{1}{3} \leqq D \leqq \frac{1}{2} \qquad \text{Eq. 18-22}$$

From this point of view, the results were quite satisfactory (see also Fig. 18-5, 18-6, 18-7 and 18-8).

In grade 5:

$m(N) = 4.92$ $\qquad$ $m_{25}(N) = 3.84$

$m(T) = 39$ $\qquad$ $m_{25}(T) = 35.5$

$m(m) = 0.99$ $\qquad$ $m_{25}(m) = 0.45$

from Eq. 18-18, 18-19 and 18-20:

$A = 4.04$

$B = 3.50$

$C = 2.58$

and from Eq. 18-21:

$D = 0.370$ which satisfies Eq. 18-22.

(It is interesting to notice that the balance of $\frac{N}{T}$ and m is very much the same as in the Budapest study where we did not use the sophisticated method of optimalizing the correlation between IA and I.Q.)

Although the IA_{raw} score is sufficient for most purposes, for easier comparison with other tests and for better understanding of the test results (especially for teachers who may be more familiar with I.Q's), the IA scores have also been converted into I.Q. equivalents (IA_{IQ}).

18.3.2 Treatment of data

The differences between the IA_{raw} scores of experimental and control classes in grades 4 and 5 separately and the differences between the grade 4 and the grade 5 classes were examined by t-tests.

18.3.3 Results

In the Sherbrooke study the IA test failed to show any significant difference between the experimental and control classes. There was, however, significant gain in IA score in the experimental classes but not the control classes between grades 4 and 5 (see Table 18-1).

As Table 18-2 shows, we received moderate but significant correlations with all the intelligence tests and subtests used in the Sherbrooke study. It is interesting to note that we received significant correlations even in the case of the Lorge-Thorndike Test, which was administered three and four years before the IA test.

The only intelligence test which was administered without a time limit in the Sherbrooke study was the Raven Test. We received significant but low

Table 18-1 EXPERIMENTAL DIFFERENCES IN THE SHERBROOKE STUDY USING IA_{raw} SCORES.

SCORES FROM THE SECRET WRITING TEST

STATISTICS / GROUP		AVERAGE		STANDARD DEVIATION		NUMBER OF SUBJECTS		F			t		
								Experimental vs Control	Grade 4 vs Grade 5: Separate for Exp. and Contr.	Grade 4 vs Grade 5: Overall	Experimental vs Control	Grade 4 vs Grade 5: Separate for Exp. and Contr.	Grade 4 vs Grade 5: Overall
4	CONTROL	2.472	1.920	3.125	2.548	20	38	5.001*	2.379*	2.019*	1.323	1.655	3.232*
	EXPERIMENTAL	1.423		1.397		18			1.121			3.807*	
5	EXPERIMENTAL	3.141	3.497	1.479	1.793	23	44	1.876			1.370		
	CONTROL	3.886		2.026		21							

*significant at the 5 per cent level.

TABLE 18-2

Correlation between IQ Tests and Subtests[1] with IA_{IQ} Scores

Tests / Grades (No of subj.)	Lorge-Thorndike	Otis-Lennon	Raven						Cattell				
			Subtests						Subtests				
			1	2	3	4	5	Total	"Series"	"Classif."	"Matrices"	"Condit."	Total
4	0.467*** (46)	0.609*** (53)	——	——	——	——	——	0.359*** (53)	——	——	——	——	0.261 (53)
5	0.403** (50)	0.542*** (55)	——	——	——	——	——	0.418** (55)	——	——	——	——	0.411** (55)
4 and 5 together	0.425*** (96)	0.563*** (108)	0.263** (108)	0.397*** (108)	0.332*** (108)	0.389*** (108)	0.368** (108)	0.389*** (108)	0.319** (108)	0.258** (108)	0.335*** (108)	0.090 (108)	0.337*** (108)

**significant at 1 percent level

***significant at 0.1 percent level

---

[1] The correlations with the subtests were based on raw scores; the others on normalized scores.

485

486

correlations between the time spent on the Raven Test and on the Secret Writing Test ($r = 0.387$, $n = 52$, $p < 0.001$ in grade 4; $r = 0.348$, $n = 55$, $p < 0.01$ in grade 5; and $r = 0.359$, $n = 107$, $p < 0.001$ in grades 4 and 5 together).

18.4 New York Study

18.4.1 Method

The method used was the same as the one in the Sherbrooke study but translated into English: "Secret Writing", a test of information acquisition and creative application (IA). A complete description of this test is given in section 18.2.1.

18.4.2 Treatment of data

IA Raw Scores were computed for each student after formula 10-16 $IA_{raw} = \frac{N^3}{T} + m$ where T refers to the time in minutes the student spent to learn the secret writing, N is the learning score (number of symbols correctly translated) and m the creative application score which is the sum of the scores which are given to each different secret writing invented by the student. For a more detailed description of the scoring system see section 18.2.1.

18.4.3 Question of interest:

Children of which school show a better performance on the learning ability test, the Fleming school students or the students of Ecole Eymard?

18.4.4 Results

As shown in table 18-3 the mean for the Fleming School students in grade 4 is 6.78 while the Eymard School students have one of 1.47 only.

Table 18-3

The differences between the Fleming School in New York and the Eymard School in Sherbrooke in the Learning Ability Test (Secret Writing), based on IA raw scores.

	School	Fleming-School (New York)			Eymard School (Sherbrooke)
GRADES	class / Statistics	red	yellow	total	total
4	mean	6.50	7.07	6.78	1.47
	Stand. deviation	4.53	3.73	4.11	1.30
	Number of subjects	21	21	42	24
	F			F = 9.97***	
	t			t =-7.73***	
5	mean	9.43	9.63	9.53	3.18
	Stand. deviation	2.91	4.10	3.48	1.39
	Number of subjects	17	16	33	27
	F			F = 6.27***	
	t			t =-9.57***	

*** significant at the .001 percent level of confidence. Variances being significantly different (F value) the t for heterogeneous variances (two-sided) was used.

The difference is significantly in favour to the Fleming School children at the .001 percent level of confidence. The same statement can be made for the results in grade 5. It is also interesting to see that the variances are significantly different in the two groups under comparison (significantly high F value) which means that among the students in the Fleming school there is a rather high variation in the performance on this test.

CHAPTER 19

CREATIVITY

Hypothesis 6. The Dienes method of mathematics learning causes greater student creativity than the traditional method.

19.1 Theoretical background

In the last 15 years, a great variety of studies on creativity has appeared all over the world. If ten years ago the question often appeared "Why can't Johnny read?" today the fashionable question is "Why isn't Johnny creative?" (Nunally, 1964).

The needs for new ideas in the area of problem solving were largely responsible for the growth of interest in this field after the second World War. The realization that creative qualities are not restricted to a gifted few but broadly distributed throughout the population is largely responsible for an increased research effort by psychologists. These views are supported by the findings of B.S. Bloom, for example, who reported in an unpublished study that 3/5 of a student population might be regarded as gifted if "belonging to the upper 10% level of performance in one or more tests of primary abilities" were used as criterion for giftedness.

Side by side with this new interest in scientific research young people inspired partly by the "hippie" philosophy or life style are to-day more prone to concentrate on detecting and developing their creative abilities, many of them refusing careers in the traditional sense within the technocratic system because it may restrict rather than develop their creative instincts. But even within the existing economic system a bright spot has appeared in the form of a change toward the realization of the basic freedom which grants the opportunity to all members of society to develop as far as possible according to their abilities. An excellent illustration of this kind of change is management reaction to General Motors' workers who went on strike because "their jobs were too dull and the ... pressure ... too intensive." (See Business Week, Sept. 9, 1972, ref. in Gordon, 1972.) A proposal to give employees a course in creative thinking was implemented, so that the workers would learn to view their jobs as a problem to be solved, solved more efficiently, and so on.

Although creativity is now in the limelight, a clear definition of the concept is not unanimously accepted by scientists. Often the term is applied to inventive people or inventive acts. It is also applied to more tangible results, such as a poem, a painting, or a scientific theory. Other explanations insist that the results must have utilitarian value. Obviously, these are not appropriate definitions in basic psychology, but consist more of a common sense view of creative people and their outputs. There is a larger class of products, psychological products which include pencil marks on paper, words, expressed ideas, even ideas merely thought of (Guilford, 1967b). To illustrate this distinction, let us cite an example from Gordon (1972):

"In the 16th century, people thought that blood flowed from heart to body, surging in and out like the tides of the sea. Harvey was familiar with this view and believed it until he closely observed a fish's heart that was still beating after the fish had been opened up. Harvey looked for a tidal flow of blood, but the action of the fish's heart reminded him instead of a pump he had seen. Harvey's creativity led to an innovative thought that was a benefit to all mankind.

Let's say a student is observing a fish's heart. He knows nothing about physiology and his professor has just explained how the heart acts like a pump. He is reminded of a swimming pool where the dirty water is pumped through the filter and back into the pool. This is an obvious connection, but he develops other connections as well. He sees how lungs and the liver act as "filters" when they cleanse the blood. Thus, through an example from his own experience the student creatively contributes to his own learning."

The learning process of the student, however, led to a private personal result which was "only" a psychological product, whereas Harvey's connection led further by communication to a public result.

If we are concerned with children's creativity and its promotion, we have to, through education, desist as far as possible from social aspects (innovation, usefulness, etc.) as part of the definition of the concept.

D.P. Ausubel obviously takes the opposite view when he writes in his article "Fostering creativity in the School": "Since creative potentialities are, by definition, sparsely distributed in the population, instances of true creativity can be anticipated no more frequently among the clientele of our schools than among any other population of human beings" (Ausubel, 1968a, p. 14). Again a statement such as "Varying degrees of creativity also exist, but the creative person differs qualitatively from individuals manifesting lesser degrees of creativity" (page 11) seems to us an elitist kind of thinking which is thwarting the hopes we place in the possibilities that education can provide the children of to-day with capacities to cope with the problems of to-morrow (see also p. 197-199 in Chapter 8, Volume I).

At the same time Ausubel admits that there are abilities (intellectual-personality functions, such as divergent thinking) which facilitate creative behavior, for us such abilities are constituents of creative behavior and a person having them in more or less of a degree (as may be assessed by appropriate tests) is a creative person, independent of the social significance his products may have.

What we are trying to do here is essentially to repeat the treatment that has been given to other concepts of common language such as "soul", "love", "intelligence", leaving them out of our terminology in psychology or giving them an operational definition which allows us to handle them in a fairly objective way.

The most objective way of assessing abilities is by test measurement. For creativity however, test development is still in progress, and there is still the problem of which constituents should fall and be reconciled under the concept creativity.

Guilford's model* of the structure of the human intellect which has made its path into psychological theory has helped our realization that there are abilities which have not as yet been covered by intelligence tests but come near to what may be referred to as creativity, especially the category of divergent thinking.

---

* This model consists of a three-dimensional figure containing 120 cells, each of which represents a different kind of intellectual ability (Guilford, 1959).

The traits which Guilford regards as belonging to the category of creativity are: Fluency of thinking, flexibility of thinking, originality, re-definition, and elaboration. Guided by the Guilford model of intellect, Torrance constructed, through adapting some of Guilford's materials and introducing new tests items, a well elaborated test of creative thinking which was used in the Sherbrooke study, also partly in the Budapest study. Before presenting these, let us have a look at how:

a) creativity is related to intelligence;

b) creativity is related to other personality traits;

c) children s creativity may be promoted in schools.

19.1.1 Creativity and intelligence

The problem of whether creativity is an aspect of ability, distinct from intelligence, has not been fully resolved, mainly because of different definitions and theoretical interpretations of both creativity and intelligence. One of the early studies done by Getzels and Jackson (1962) led to what is called the threshold theory, which asserts that beyond a certain I.Q. (the suggested cut-off is 120), there is little or no relationship between intelligence and creativity and that being more intelligent is no guarantee of correspondingly greater powers of creativity. While Wallach and Kogan (1965) criticize the tests of "creativity" often correlate as well or better with measures of I.Q. than they correlate among themselves, Guilford (1967b) argued that creative talent may be so complex that current measures (such as divergent thinking) may well tap quite unique aspects of the construct, and as such can not be expected to display high intercorrelations. The creativity-intelligence question may approach its solution only after multivariate statistical techniques take the place of simple correlation studies.

19.1.2 Personality correlates of creativity*

Considerable research has been conducted on the personality characteristics of persons who have been rated by competent judges to be creative in such areas as art, architecture, literature, and science. On the cognitive side, creative individuals tend to be original, perceptive, insightful, independent in judgment, open to new experience (especially from within], sceptical and verbally facile. They are flexible, open-minded and tolerant of ambiguity; have wide-ranging interests; prefer complexity; and are less interested in small details and in the practical and concrete than in theoretical ideas and symbolic transformations (Barron, 1963; Drevdahl, 1956; Drevdahl & Cattell, 1958; MacKinnon, 1962). In general they delight in paradoxes and in reconciling opposites. From a motivational standpoint they are ambitious, achievement-oriented, dominant, and have a sense of destiny about themselves. They tend to be emotionally and aesthetically sensitive. Their self-image abounds in such traits as inventiveness, determination, industry, independence, individualism, and enthusiasm. On the whole, they exhibit higher ego strength and self-acceptance, more introspectiveness, and greater femininity than non-creative individuals. In their relation with others they are unconventional, rebellious, disorderly, self-centered exhibitionistic and prone to retreat to the role of observer (Barron , 1963; Drevdahl, 1956, Drevdahl & Cattell, 1958; Hammer, 1961; MacKinnon, 1960, 1961, 1962).

---

* After Ausubel 1968a.

As measured by their tests of creativity, Wallach and Kogan (1965) found that their high creatives tended to be broad rather than narrow categorizers, to be tolerant of an unconventional type of hypothesizing about the world, and to be responsive to affective aspects of the environment. The high creative/high intelligent groups were high in self-confidence and self-esteem and low in defensiveness; enjoyed a high sociometric status; actively sought the companionship of others; and exhibited a high attention span and ability to concentrate, but at the same time tended to display more than their fair share of attention seeking and disruptive behavior. On the other hand, high creatives who were low in intelligence exhibited the opposite set of characteristics except for attention seeking, disruptive behavior in the classroom. Anxiety level was middling for the high creative group ; when it was either very high or very low, it appeared to depress creativity. This finding suggests that either a moderate degree of anxiety is necessary to generate creative behavior, or that the expression of creativity is productive of some anxiety.

19.1.3 Fostering creativity in the school

There are two methods of fostering creativity in school: On the one hand we can try to select the most creativity-gifted individuals and see to it that they get the opportunity "to let their genius out"; on the other hand, we can attempt to change the curriculum, the teaching methods and teachers' attitude in order to train every child in thinking and approaching problems in a more creative manner.

Ausubel (1968a) with his comprehension of creativity tends towards the first way in proposing to help only the few: "Creativity is a rare gift. The school can help only in actualizing its expression in those who already possess the potentiality" (1968, p.14). We, on the other hand, propose the second and more democratic way. In accord with Kurt Danziger (1968) we feel that it is time to overcome the charismatic view of creativity, which belongs to the Renaissance model of the creative individual, the artist. That is certainly not the only field where creative behaviour takes place and creativity is not an awesome and inexplainable property belonging to a few chosen souls, but rather to some (though yet indeterminable) extent a situational product. We do not yet know which part of the role genetic factors play but several studies show (Getzels and Jackson 1962, Weisberg and Springer 1961, and others) that certain motivational and environmental factors are associated with high creativity.

Next to the family the school is the most important environmental factor in this case. The question how far it is realistic to expect the promotion of creativity from an institution such as a school, whose primary function is the transmission rather than the change of the existing culture, is too involved to be discussed in detail here (see Danziger, 1968; Henry, 1963).

The fact, however, that creativity can be trained has been demonstrated by research. As far as originality is concerned, subjects can be trained to respond with more unusual associations, to generate more novel ideas, or to suggest more usual uses (Mednick 1962, Freedman 1965, Maltzman 1960), and in some instances transfer to related kinds of problems does take place. Rusch et al., (1965) compared an experimental group taught "creatively" for the school year and a control group that was taught in the regular manner.

The experimental group did significantly better on five out of seven divergent production tests. While Brown (1964) was concerned with the classroom milieu, Wilson, Stuckey and Langevin (1972) compared two open plan schools (which were not graded and emphasizing a discovery approach to learning) against two control schools and could find higher scores on some of Torrance's creativity tests in the former.

Torrance et al. (1960) taught 10 teachers some principles of how to teach creatively, and the result was measurable on student performance in creativity tests. His direct approach of the interested and open-minded teacher may be one useful way of nurtering creativity in schools. Let us therefore recall some of his suggested principles (Torrance, 1969):

Provide materials that develop imagination
Provide materials that enrich imagery
Permit time for thinking and day dreaming
Encourage children to record their ideas
Give children's ideas some concrete embodiment
Accept and use the tendency to take a different look
Prize rather than punish true individually
Be cautious in editing children's productions
Encourage children to play with words
Use of incompleteness and openness to create intrinsic motivation
Produce things and use them.

In a program such as that of Dr. Dienes creativity is fostered all the while. Torrance's principles are both built into the theoretical framework of the mathematics curriculum and implemented in the classroom from day to day.

19.2 Budapest study

19.2.1 Method

The students' creativity has been investigated using three parts of Torrance's Tests of Creative Thinking (1962, 1966):

a. The "Circles" test (a picture completion test) suggesting additions to circles printed on a page or drawings containing the circles as their main parts.

b. The "Dog" test (a "product improvement" test) suggesting changes to a stuffed toy dog so as to make it "better, nicer, more interesting".

c. The Unusual Uses test suggesting the invention of uncommon ways of using objects familiar to children. (Instead of "coins" and "cardboard boxes" used by Torrance, we adopted "newspaper", "button", and "key" from amongst those suggested by Wallach and Kogan (1965)).

The factors discerned were fluency, flexibility and originality (except in the Picture Completion test, for which only fluency and originality were scored because of the difficulty of grouping the answers), in the traditional sense (Torrance, 1966) but with our own criteria:

a. The measurement of fluency was the number of answers.

b. The measurement of flexibility was the number of groups occurring in the subject's answers (to group the answers we took into consideration both Torrance (1966) and our earlier study (Barkóczy-Klein, 1968)).

c. For originality we gave a score of

0 for an answer which was given by 20% or more of the students;

1 for an answer which was given by 5% or more but less than 20% of the students;

2 for an answer which was given by 2% or more but less than 5% of the students;

3 for an answer which was given by less than 2% of the students.

We considered, as a student's originality score, the sum of the originality scores of his answers.

The examination of creativity was carried out as an individual test in grade 3, and as a group test in grades 4 and 5.

19.2.2 Treatment of data

To examine the differences in fluency between experimental and control classes we used the two-sample t-test if the F ratios were not significantly different, and the d-test if they were, in both cases supposing that the fluency scores followed a normal distribution.

Because we could not assume a normal distribution of the flexibility and originality scores, we used the chi-square test for flexibility, while for the more widely distributed originality scores we used the Mann-Whitney test.

19.2.3 Expectation

Higher fluency, flexibility, and originality scores in experimental classes than in control classes would show greater creativity, thus supporting hypothesis 6.

19.2.4 Results

The results of the examinations of student creativity are shown in Table 19-1. Let us look at these results test by test and factor by factor.

Table 19-1

Summary of the results of the examinations of student creativity in the Budapest Study.

Class		Picture Completion ("circles")		Product Improvement ("Dog")			Unusual Uses ("Button", "Newspaper", "Key")		
		Fluency	Originality	Fluency	Flexibility	Originality	Fluency	Flexibility	Originality
3	Experimental	6.5 3.6	7.4 6.2	12.5 4.7	4.8 1.7	9.9 5.8	15.3 1.7	9.7 3.8	23.7 10.6
	Control	4.6 2.8	4.0 4.5	4.9 5.6	4.6 1.8	8.7 4.9	17.0 3.3	10.3 3.5	21.4 10.4
	Difference	$t = 2.13^*$	Mann-Whitney *	$t = 4.4^*$	-	Mann-Whitney	$d = 2.15^{\boxed{*}}$	-	Mann-Whitney
4	Experimental	5.6 3.6	5.4 3.6	6.3 2.1	4.6 1.3	7.0 2.5	13.1 5.8	11.4 2.3	15.7 8.6
	Control	8.2 2.3	6.4 4.0	6.5 2.6	3.0 1.5	4.5 3.2	14.0 5.9	9.3 3.3	12.6 7.2
	Difference	$t = 3.2^{\boxed{*}}$	Mann-Whitney	$t = 0.35$	$\chi^2 = 2.9$	Mann-Whitney *	$t = 0.53$	$\chi^2 = 2.58$	Mann-Whitney
5	Experimental	9.7 3.5	9.5 5.6	9.2 4.4	4.7 2.2	9.4 3.2	19.4 8.2	12.6 4.3	28.1 2.7
	Control	9.6 3.4	8.8 5.9	5.9 4.1	4.0 2.0	5.4 4.2	13.6 4.8	9.4 3.1	17.0 6.9
	Difference	$t = 0.092$	Mann-Whitney	$t = 2.49^*$	-	Mann-Whitney *	$d = 2.52^*$	$\chi^2 = 4.4^*$	Mann-Whitney *

5.6	arithmetic mean
5.6	standard deviation
$\boxed{*}$	significant difference favouring the control class
*	significant difference favouring the experimental class

"CIRCLES" (PICTURE COMPLETION)

Fluency

The results in this factor were confusing: while in grade 3 the experimental class was significantly better, grade 4 gave the opposite result, and in grade 5 there was no significant difference.

There are at least two considerations which could interfere with using the fluency scores obtained from this test as a parameter of creativity:

i. The task requires subjects to make the circle a part of the drawing, so those answers in which the drawing is simply done within the circle are not counted. This means that the subjects have to change their reference level, in this case from the blank page to the circle. This is quite difficult until about age 8, but later it becomes quite easy. As the experimental mathematics teaching gives many opportunities for the students to change their reference level in different activities, it could cause a significant difference in the early years which might then be equalized simply by the natural mental development of the children.

This could explain the disappearance of the differences in later years but not the significantly better result in the control class in grade 4.

ii. The result for this factor in this test is very dependent on the drawing ability of the children, a fact which could cause a significant but from our point of view uninteresting difference between the experimental and control classes.

Originality

Both explanations are partially supported by the originality results:

i. The significant difference in grade 3 does disappear as the subjects become older.

ii. There is no significant difference in grade 4 in originality, which is normally considered a better parameter of creativity.

"DOG" (PRODUCT IMPROVEMENT)

Fluency

There were significantly better results in the experimental classes in grades 3 and 5 but no significant difference in grade 4. Apart from the possibility of differences between the classes in different grades, the latter result can be explained by the fact that, while in grade 3 the individual testing method was used, in grades 4 and 5 the testing was in groups. This could affect the results mainly in two ways:

i. The experimenter's presence may have the effect of stimulating answers.

ii. While in the case of individual testing the experimenter writes the answers, in the case of the group test the answers are written by the individual himself, a process which can be much slower, especially in the lower grades.

In fact, in grade 4 there was a much lower average fluency score in the experimental class than in grades 3 or 5.

Flexibility

There were no significant results in terms of this factor.

Originality

The above-mentioned change from individual to group testing did not affect the results in originality. There were significantly better results in the grade 4 and 5 experimental classes and the results in this factor clearly suggest that the differences between the experimental and control classes in the students' originality increase with the time the student has spent in school.

UNUSUAL USES

This factor in this test shows a quite continuous change with time: while in grade 3 the control class showed a significantly better result, in grade 4 there was no significant difference and already in grade 5 the experimental class had become significantly better.

Flexibility

The only significant difference in flexibility in this test was in grade 5, in which the differences in all three parameters were significant. In grade 3 the results of the control class were slightly better and in grade 4 the results in the experimental classes were slightly better. We could suppose a continuous improvement in the experimental class in this factor as well.

Originality

As above, the advantage of the control classes, not significant in grades 3 and 4, became significant in grade 5.

If we look now at the results of each factor, we can find 6 significant differences in fluency of which 2 were in the opposite direction from our hypothesis. In flexibility there was only 1 significant difference, while in originality there were 4, all supporting our hypothesis. So, probably because of the uncertainty in the scoring, flexibility proved to be the least sensitive factor. Fluency, a more quantitative measure, proved to be the most sensitive, while originality, which is more qualitative, showed the differences unambiguously.

SUMMARY

Considering the differences by grades, we find that in grade 3, three significant results support our hypothesis while one opposes it. In grade 4 there was only one significant difference supporting and one opposing our hypothesis, while in grade 5, five differences supported it. So, although it is clear that the lead of the experimental classes is greatest in grade 5, to show a continuous development in creativity (especially because of the problems in grade 4) a longitudinal study would be necessary.

19.3 Sherbrooke Study

19.3.1 Method

As in the Budapest study, measuring creativity is based on Torrance Tests of Creative Thinking (1966, 1968, 1970; Tests de Pensée Créative, 1972). All subtests of Form A (French version) were administered to the children of all three grades in both the experimental and the control school. However, because of the immense work involved in scoring the tests (new categories and standards of frequency had to be established) from the verbal part only one subtest could be scored. For this only grades 4 and 5 were considered because group testing turned out to deliver no reliable results in grade 3.

SUBTESTS

Evaluating the children's creativity in the Sherbrooke study the following subtests have been used:

Picture Construction:

This test requires subjects to create a picture, using standardised shapes of coloured paper (here a green ellipse) as the base stimulus, and making the paper shapes an integral part of the picture. The instructions require the subject to make up as highly elaborated and unusual a picture as they can. They are also asked to make up titles for their pictures.

Picture Completion:

Subjects are required to finish off ten incompleted figures and make up an interesting title.

Parallel Lines:

In these tests subjects are presented with a page on which, in one form, thirty parallel lines have been drawn. The pair of lines are not elaborated on in any way, so that the subject is confronted with a page full of lines. The instructions require him to draw as many interesting and unusual things as he can think of in which the lines form an integral part. Underneath he is asked to write the name of the object. Common responses include a house, a ladder, and so on.

Product improvement:

In this test subjects are shown a common toy (here a stuffed elephant six inches tall) and then are asked to list the "cleverest, most interesting and unusual ways" for changing this toy elephant so that children will have more fun playing with it.

SCORING

Figural part

In scoring the figural part of the Torrance Test of Creative Thinking, we mainly followed the suggestions made by Torrance in the "Directions Manual and Scoring Guide" (revision January 1970).

The main deviations were:

1. Elaboration was scored only in Activity 1.
2. Categories on which scoring of flexibility is based were slightly changed. Some were left out, some added, and some divided (see Appendix IV).

3. For originality three different scores were used O_f and O_{add} and the sum of O_f and O_{add}. O_f is based on frequency of the response, as is the Torrance originality score, but the frequencies were newly determined in accordance with our population (166 subjects from grades 3, 4 and 5 in an elementary school). O_{add} was taken as an additional score, because so many drawings (especially in Activity 2) were very complex and "fantastic" that the experimenter would lose information by scoring these only in regard to frequency (see also Yamamoto, 1964). An elaboration score would not have covered the amount of creative strength, because in scoring elaboration no difference is made between usual and unusual details.

The score O_{add} ranged from 0 to 3:

One point was given a) when another main concept was added (e.g., a tree with an animal in its trunk; mountains with people climbing them), or b) if an activity was expressed (a girl swimming; flowers in the wind).

Two points were given a) when the activity consisted of more than a movement, i.e. a relationship or an action between two persons or things (the marriage of two birds; a cowboy throwing a lasso over a cow's neck), or b) when a special meaning was given to the objects (x is the friend of y; "tourbillon de la mort"; "les montagnes amoureuses").

Three points were given in the case where a whole story or a joke was told (St. Peter fell through the clouds because he had eaten too much iron). This score was given only a few times in Activities 1 and 2.

4. Fluency, flexibility and originality scores were not only used in their raw form but were also transformed into normalized scores.

5. A creativity score was defined as the average of the normalized fluency, flexibility and originality scores.

Verbal part

For Scoring the Product Improvement Subtest from the Verbal Part of the Torrance Test of Creative Thinking we defined the categories for our own population (see Appendix IV).

Fluency, flexibility and originality (originality 1 with our notion) were defined from these categories according to Torrance with the number of adequate responses, the number of categories being used, and the sum of the originality scores of the responses. An originality 1 score was 0, 1, 2 or 3 depending on the frequency with which the response occurred.

An additional originality score (originality 2) was defined as originality 1 divided by the fluency score.

Four kinds of combined scores were defined after Barkóczi and Klein, 1968, as follows:

Given a certain question and a certain population, for every subcategory, by Eq. 19-1 an originality score $r(c_{mn})$ can be computed.

$$r(c_{mn}) = 1 - \left(\frac{I(c_m) + I(c_{mn})}{T}\right)^b \qquad \text{Eq. 19-1}$$

where c_m stands for categories

c_{mn} for subcategories in category c_m

$I(c_m)$ is the number of answers in category c_m $(m = 1, 2 \ldots M)$

$I(c_{mn})$ is the number of answers in subcategory c_{mn} $(n = 1, 2 \ldots N_m)$

T is the number of answers to the question in this population, and

b is an exponent defined so that r would vary practically between 0 and 1.

For each subject j, three creativity scores were computed (see Eq. 19-2, 19-3 and 19-4):

$$A(j) = \sum_{m=1}^{M} \sum_{n=1}^{N_m} \sum_{\ell=1}^{L(j)} R(c_{mn}, \ell, j) \qquad \text{Eq. 19-2}$$

$$\text{where } R(c_{mn}, \ell, j) \begin{cases} = r_{mn}), \text{ if subject } j \text{ gave his } \ell^{th} \text{ answer in subcategory } c_{mn} \\ = 0, \text{ if not.} \end{cases}$$

and $L(j)$ stands for the number of answers subject j gave.

In addition to the originality, $A(j)$ takes into consideration the fluency.

$$B(j) = \sum_{m=1}^{M} T(c_m, j) \qquad \text{Eq. 19-3}$$

$$\text{where } T(c_m, j) = \max \left[R(c_{mn}, \ell, j) \right]$$

B(j) in addition to originality takes into consideration flexibility as well; and

$$C(j) = \frac{A(j)}{L(j)} \qquad \text{Eq. 19-4}$$

which is presumed to be a pure measurement of the originality of the subject j. The fourth way of combining fluency, flexibility and originality was the following:[*]

$$k(j) = \sum_{m} \frac{T - I(c_m)}{T} \sqrt{i(c_m, j)} \qquad \text{Eq. 19-5}$$

where $I(c_m)$ is the number of answers in category c_m,

$i(c_m, j)$ is the number of answers in category c_m, given by subject j.

19.3.2 Treatment of data

The frequency distributions of the different factors (fluency, flexibility, originality and elaboration) in the different activities (activities 1, 2, 3 separately and all combined) of the Figural Part of the Torrance Test of Creativity were constructed for the three age groups separately and combined.

Two kinds of statistic have been used on the scores:

i) Correlation coefficients have been calculated to see the connection between the different parts, factors and scoring methods of the Torrance Test.

ii) t-tests have been made to see the differences between boys and girls and children in the experimental and control classes.

In the four cases where significant differences between experimental and control groups have been found the results were also graphed.

---

[*] This method was suggested by B. Connes.

19.3.3 Results

FREQUENCY DISTRIBUTIONS - EFFECT OF AGE

Figural Part

Figures[*] 19-1 through 19-11 show the frequency distributions of the different factors in the figural part of the Test. As can be seen, the shapes of the distributions are very much alike in the different age groups, although the scores tend to grow from grade 3 to grade 5. For example, in activities 1, 2 and 3 combined, and also for the experimental and control classes taken together we obtained:

Fluency: $\bar{X}_{\text{grade } 3} = 19.27$, $\bar{X}_{\text{grade } 5} = 24.16$

Flexibility: $\bar{X}_{\text{grade } 3} = 15.24$, $\bar{X}_{\text{grade } 5} = 18.32$

Originality($O_f + O_{add}$): $\bar{X}_{\text{grade } 3} = 28.39$, $\bar{X}_{\text{grade } 5} = 44.9$

Most of the frequency distributions are quite symmetrical (with the exception of fluency in activity 2, flexibility in activity 2 and originality ($O_f + O_{add}$) in activity 1 and activity 3). Because of this, the normalization of the scores did not result in big changes (see later).

Verbal Part

As can be seen from Table 19-10 all except the originality scores did show a growing tendency with age:

A (originality and fluency): $\bar{X}_{\text{grade } 4} = 4.77$, $\bar{X}_{\text{grade } 5} = 5.28$

B (originality and flexibility): $\bar{X}_{\text{grade } 4} = 2.51$, $\bar{X}_{\text{grade } 5} = 2.96$

K_v (creativity): $\bar{X}_{\text{grade } 4} = 5.67$, $\bar{X}_{\text{grade } 5} = 6.77$

Fluency: $\bar{X}_{\text{grade } 4} = 11.97$, $\bar{X}_{\text{grade } 5} = 12.73$

Flexibility: $\bar{X}_{\text{grade } 4} = 4.21$, $\bar{X}_{\text{grade } 5} = 5.27$

---

[*] All these Figures will be found at the end of this chapter.

On the other hand, the results in C (originality), in originality 1 and originality 2 slightly declined from grade 4 to grade 5.

CORRELATION

Figural Part

1. Looking at the four tables (Tables 19-2 through 19-5) which contain the intercorrelations of the Figural Part of the Torrance Test of Creative Thinking, the first thing that can be noticed is that there are no big differences between the correlation coefficients found in the different grades. Because of this we shall hereafter refer only to Table 19-5 which contains data about all the grades combined.

2. Let us first consider the intercorrelation between the different kinds of originality score:

i) Looking at the correlations between scores in the same scoring system (O_f, O_{add} or $O_f + O_{add}$) but in different activities we found significant but low correlations ranging from .14 to .39.

ii) The intercorrelations between the three different kinds of originality score (O_f, O_{add} and $O_f + O_{add}$) in the same activity show that the Torrance kind of score and the additional score are independent (r_{O_f}/O_{add} = -.6 in activity 1, -.10 in activity 2 and -.5 in activity 3). The final score is mainly a function of the Torrance kind of score ($r_{O_f} / [O_f + O_{add}]$) = .80 in activity 1, .84 in activity 2 and 1.00 in activity 3), but in all activities except activity 3 it is also affected by the additional score ($r_{O_{add}} / [O_f + O_{add}]$ = .55 in activity 1 and .46 in activity 2). In what follows we shall only take into consideration the final originality score ($O_f + O_{add}$).

3. Let us now consider the intercorrelations between the four main factors of fluency, flexibility, originality and elaboration.

Table 19-2 The intercorrelations of raw scores derived from the Torrance Test of Creative Thinking (French Form; Torrance, 1972) Figural Part in grade 3 (n = 58)

FACTORS			FLUENCY		FLEXIBILITY		ORIGINALITY								
							After Torrance $(\bar{O}_f)$			Additional $(\bar{O}_{add})$			Final score $[(\bar{O}_f) + (\bar{O}_{add})]$		
		Activity	2	3	2	3	1	2	3	1	2	3	1	2	3
Fluency		3	.33**												
Flexibility		2	.86***	.29*											
		3	.25	.87***	.24										
Originality	after Torrance	1	.14	.01	.08	.13									
		2	.68***	.05	.65***	.05	.08								
		3	.32*	.90***	.28*	.88***	.11	.07							
	Additional score	1	.14	.00	.14	-.06	-.16	.16	-.03						
		2	-.05	.10	-.06	.03	.27*	-.04	.12	.03					
		3	-.28*	-.13	-.24	-.07	.11	-.19	-.16	.12	.39**				
	Final score	1	.20	.02	.15	.09	.85***	.16	.09	.39**	.27	.16			
		2	.61***	.09	.58***	.06	.19	.91	.11	.16	.38**	-.01	.26*		
		3	.29*	.89***	.26*	.88***	.12	.05	1.00***	-.02	.16	-.07	.10	.11	
Elaboration		1	.13	-.02	.00	.01	.30*	.10	.05	.43***	.30**	.23	.51***	.22	.08

* significant at 5 percent level
** significant at 1 percent level
*** significant at 0.1 percent level

Table 19-3 The intercorrelations of raw scores derived from the Torrance Test of Creative Thinking (French Form: Torrance, 1972) Figural Part in grade 4 (n = 52)

FACTORS			FLUENCY		FLEXIBILITY		ORIGINALITY								
							After Torrance (O_f)			Additional score (O_{add})			Final score $[(O_f) + (O_{add})]$		
		Activity	2	3	2	3	1	2	3	1	2	3	1	2	3
Fluency		3	.62***												
Flexibility		2	.93***	.60***											
		3	.53***	.74***	.49***										
Originality	after Torrance	1	.11	.25	.17	.30*									
		2	.68***	.61***	.74***	.44**	.24								
		3	.57***	.87***	.60***	.81***	.25	.58***							
	Additional score	1	-.13	-.07	-.08	.12	.04	-.06	-.01						
		2	-.01	-.08	-.03	-.11	-.03	-.03	-.11	.32*					
		3	-.20	-.19	-.18	-.16	.02	-.15	-.18	.12	.37**				
	Final score	1	.03	.17	.09	.32*	.85***	.16	.20	.55***	.14	.07			
		2	.58***	.49***	.63***	.33*	.19	.85***	.44**	.11	.50***	.06	.22		
		3	.56***	.87***	.59***	.80***	.25	.57***	1.00***	.00	-.08	-.09	.21	.45***	
Elaboration		1	.01	.01	.10	.06	.13	.18	.16	.18	.40**	.27	.21	.37	.19

* significant at 5 percent level
** significant at 1 percent level
*** significant at 0.1 percent level

Table 19-4 The intercorrelations of raw scores derived from the Torrance Test of Creative Thinking (French Form; Torrance, 1972) Figural Part in grade 5 (n = 55)

FACTORS			FLUENCY		FLEXIBILITY		ORIGINALITY								
							After Torrance (O_f)			Additional score (O_{add})			Final score $[(O_f) + (O_{add})]$		
		Activity	2	3	2	3	1	2	3	1	2	3	1	2	3
Fluency		3	.42***												
Flexibility		2	.70***	.54***											
		3	.37**	.89***	.51***										
Originality	after Torrance	1	.13	.19	.22	.19									
		2	.44***	.35**	.54***	.37**	.09								
		3	.42***	.93***	.52***	.87***	.19	.26							
	Additional score	1	-.14	-.17	-.23'	-.19	-.13	-.31	-.11						
		2	-.15	-.21	-.17	-.20	-.00	-.28*	-.12	.59***					
		3	-.34**	-.08	-.26	-.10	-.14	-.21	-.07	.35**	.27*				
	Final score	1	.00	.03	.00	.01	.69***	-.16	.07	.62***	.43***	.14			
		2	.29	.16	.36**	.19	.08	.70***	.15	.15	.49***	.01	.17		
		3	.39**	.92***	.50***	.86***	.18	.24	1.00***	-.07	-.09	.02	.09	.15	
Elaboration		1	-.15	-.08	-.29*	-.11	-.04	-.19	-.11	.42***	.38**	.33*	.27*	.11	-.08

* significant at 5 percent level
** significant at 1 percent level
*** significant at 0.1 percent level

Table 19-5 The intercorrelations of raw scores derived from the Torrance Test of Creative Thinking (French Form; Torrance, 1972) Figural Part in grade 3, 4 and 5 together (n = 165)

FACTORS			FLUENCY		FLEXIBILITY		ORIGINALITY								
							After Torrance (O_f)			Additional score (O_{add})			Final score ($O_f + O_{add}$)		
		Activity	2	3	2	3	1	2	3	1	2	3	1	2	3
Fluency		3	.44***												
Flexibility		2	.84***	.46***											
		3	.36***	.86***	.38***										
Originality	after Torrance	1	.13	.17	.15	.23**									
		2	.59***	.30***	.65***	.25***	.14								
		3	.42***	.90***	.44***	.86***	.22**	.28***							
	Additional score	1	-.03	-.04	-.06	-.02	-.06	-.07	.00						
		2	-.05	-.02	-.07	-.03	.10	-.10	.03	.39***					
		3	-.25***	-.07	-.21*	-.05	.00	-.16*	-.05	.26***	.35***				
	Final score	1	.09	.12	.09	.18*	.80***	.07	.18	.55***	.32***	.16*			
		2	.50***	.26***	.54***	.21*	.18*	.84***	.26***	.15	.46***	.04	.24**		
		3	.40***	.90***	.42***	.85***	.22**	.26***	1.00***	.03	.06	.04	.20*	.26***	
Elaboration		1	-.01	-.03	-.05	-.02	.14	.06	.01	.33***	.34***	.27***	.32***	.24**	.04

* significant at 5 percent level
** significant at 1 percent level
*** significant at 0.1 percent level

i) Looking at the correlations between the scores in the same factor (fluency, flexibility or originality) but in different activities we find significant but low correlations (from .20 to .44). It is interesting to note that the lowest inter-activity correlation coefficients are those in originality and the highest those in fluency.

ii) Looking at the correlation coefficients between scores in the same activity but in different factors we can find a very high correlation between fluency, flexibility and originality ($r_{flu/flex} = .86$, $r_{flu/orig.} = .90$, $r_{flex/orig} = .85$) in activity 3, and between fluency and flexibility in activity 2 (.84). The correlation between fluency and originality as well as between flexibility and originality is .5 in activity 2. Originality and elaboration have a correlation coefficient of .3 in activity 1.

iii) On the right side of Table 19-6 can be found the intercorrelations between the three main factors and the creativity score defined as the average of the normalized fluency, flexibility and originality scores in the three activities taken together. It can be seen that these correlations are very high (between .83 and .95).

Verbal Part

Although there are some differences between grades 4 and 5 in the following we shall refer only to results derived from the combined grade 4 and 5 population (see Table 19-6).

1. Let us consider first the three factors fluency, flexibility and originality 1, as scored traditionally (after Torrance). The three correlation coefficients are .46 (fluency-flexibility), .53 (fluency - originality 1), and .65 (flexibility - originality 1) all about in the order and range that could be expected from the literature.

The intercorrelations between the different scores of the verbal (based on 1 test) and the figural part (based on 3 tests) in the Torrance Tests of Creative Thinking.

Table 19 - 6

		Verbal Part								Figural Part		
		A	C	B	Kv	FLU	FLEX	ORIG.1	ORIG.2	FLU	FLEX	ORIG (Of+Odd)
Grade 4												
Verbal Part	C	0.44										
	B	0.76	0.53									
	Ky	0.90	0.25	0.88								
	FLU	0.84	-0.07	0.50	0.83							
	FLEX	0.67	0.27	0.95	0.89	0.50						
	ORIG.1	0.76	0.30	0.68	0.74	0.63	0.64					
	ORIG.2	0.06	0.40	0.25	0.03	-0.16	0.19	0.56				
Figural Part	FLU	-0.01	0.08	0.17	0.06	-0.12	0.20	0.21	0.23			
	FLEX	-0.08	-0.00	0.14	0.03	-0.15	0.16	0.10	0.09	0.94		
	ORIG (Of+Odd)	-0.02	0.07	0.17	0.04	-0.15	0.18	0.13	0.15	0.83	0.83	
	CREATIVITY	-0.02	0.10	0.14	0.02	-0.16	0.15	0.16	0.16	0.92	0.90	0.94
Grade 5												
Verbal Part	C	0.35										
	B	0.76	0.44									
	Ky	0.89	0.14	0.88								
	FLU	0.84	-0.14	0.51	0.83							
	FLEX	0.68	0.25	0.95	0.89	0.52						
	ORIG.1	0.76	0.26	0.68	0.73	0.63	0.64					
	ORIG.2	-0.02	0.51	0.15	-0.09	-0.24	0.05	0.44				
Figural Part	FLU	-0.01	0.09	0.17	0.06	-0.12	0.19	0.21	0.20			
	FLEX	-0.08	0.02	0.14	0.02	-0.15	0.16	0.10	0.11	0.94		
	ORIG (Of+Odd)	-0.01	0.06	0.17	0.05	-0.14	0.19	0.13	0.12	0.83	0.83	
	CREATIVITY	-0.02	0.10	0.14	0.02	-0.16	0.15	0.16	0.15	0.92	0.90	0.94
Grade 4+5												
Verbal Part	C	0.33										
	B	0.83	0.41									
	Kv	0.41	-0.01	0.39								
	FLU	0.69	-0.30	0.44	0.37							
	FLEX	0.78	0.26	0.96	0.42	0.46						
	ORIG.1	0.74	0.22	0.70	0.22	0.53	0.65					
	ORIG.2	0.07	0.55	0.21	-0.15	-0.23	0.12	0.47				
Figural Part	FLU	-0.04	0.02	0.10	0.09	-0.07	0.13	-0.02	0.05			
	FLEX	-0.06	-0.01	0.09	0.09	-0.08	0.12	-0.03	0.00	0.94		
	ORIG (Of+Odd)	-0.07	-0.02	0.08	0.07	-0.08	0.11	-0.05	-0.01	0.85	0.87	
	CREATIVITY	-0.04	0.01	0.09	0.10	-0.07	0.12	-0.02	0.00	0.92	0.92	0.95

2. Originality 2 (originality 1 divided by fluency) has a .47 correlation with originality 1, and a .55 correlation with C, the "pure" originality measure constructed in the same way as the measure for originality 2.

3. Looking at the "new" measurements we can make the following observations:

   i) A (originality and fluency) is independent of originality 2 (probably because A is so much affected by fluency), has a low correlation with C (.33) and with K_v (.41), but a fairly high correlation with the other measures (.83 with B, .69 with fluency, .78 with flexibility, .74 with originality 1).

   ii) B (which is supposed to be a combination of originality and flexibility) has a .96 correlation with flexibility, showing that it is hardly affected by originality.

   iii) C (originality), as has been said before, is nearest to the other "pure" originality measure (originality 2). It has a negative correlation with fluency (-.30) and is slightly connected with flexibility through B (.44).

   iv) As regards the new creativity measurement K_v we get a better picture by looking at grades 4 and 5 separately. It can be seen that with the exception of the two "pure" originality measures (C and originality 2) it correlates highly with all the other measurements (between .73 and .90).

Verbal and Figural Part

There is practically no connection between the Verbal and Figural Parts of the Test.

DIFFERENCES

Effect of Sex

We were looking for sex differences only in the Verbal Part of the Torrance Test (see Table* 19-7). Only a few differences were noticeable, and these always favoured the girls.

Grade 4 girls gave a better performance than grade 4 boys in:

B (originality and flexibility): $\bar{X}_{boys} = 2.0$, $\bar{X}_{girls} = 3.4$, $t = 2.68$, $p < 0.5$

C (originality): $\bar{X}_{boys} = 0.41$, $\bar{X}_{girls} = 0.59$, $t = 2.57$, $p < 0.5$.

Flexibility: $\bar{X}_{boys} = 3.5$, $\bar{X}_{girls} = 5.4$, $t = 2.39$, $p < 0.5$.

Again, grade 5 girls were superior to grade 5 boys in C.

Experimental - Control Differences

We were looking for these differences in both the raw (Table 19-8) and normalized (Table 19-9) scores for the Figural Part, and in the raw (Table 19-10) scores for the Verbal Part of the Torrance Test. Table 19-11 gives a summary of the significant differences found in all the creativity measurements that had been used.

The first thing to notice is that the Verbal Part of the Test (with all the new and fairly complicated scores) failed to reveal significant differences, except in the case of grade 4 fluency where the control group proved to be the better one.

The Figural Part of the Test on the other hand clearly gives a more favourable picture of the creativity of the experimental classes than of the control classes.

---

* Tables 19-7 to 19-11 will be found at the end of this chapter.

In what follows we count the number of significant differences regardless of whether they are to be found in the raw or the normalized scores (Table 19-11).

If we look at the results for each factor we can find 8 out of 12 possible significant differences for fluency. Each of these favours the experimental classes (see Fig. 19-12 and 19-13).

For flexibility there was only 1 significant difference (in grade 5, favouring the experimental class), while for originality ($0_f + 0_{add}$) there were 7 out of 16 possible ones, all supporting our hypothesis (see Fig. 19-14 and 19-15).

Elaboration was scored only in Activity 1 and gave a significantly higher score for the control class in grade 4 and for the experimental class in grade 5.

It is interesting to note that very much as in the Budapest Study (see 19.2.4) flexibility proved to be the least and fluency the most sensitive factor, whereas originality showed the differences unambiguously, even after the verbal part had been taken into consideration as well.

Looking at the differences by grades it can be seen that in grade 3 there is no significant difference between the experimental and control class in the main factors; in grade 4 the significant differences can be found (in Activity 1) only in fluency, the most quantitative measure, in favour of the experimental class, and in elaboration in favour of the control class; in grade 5 the experimental class gave more favourable results in fluency (except in Activity 2), elaboration and originality ($O_f + O_{add}$) in all activities and consequently in scores for all the three activities combined.

The creativity score, calculated from the normalized fluency, flexibility and originality scores in all three activities is significantly higher in the experimental classes (grades 3, 4 and 5 together) mainly because of the significant differences in grade 5.

All this suggests the conclusion that just as in the Budapest Study (see 19.2.4) the difference between the creativity of the children in the experimental and control classes seems to grow with the years that are spent in different educational environments.

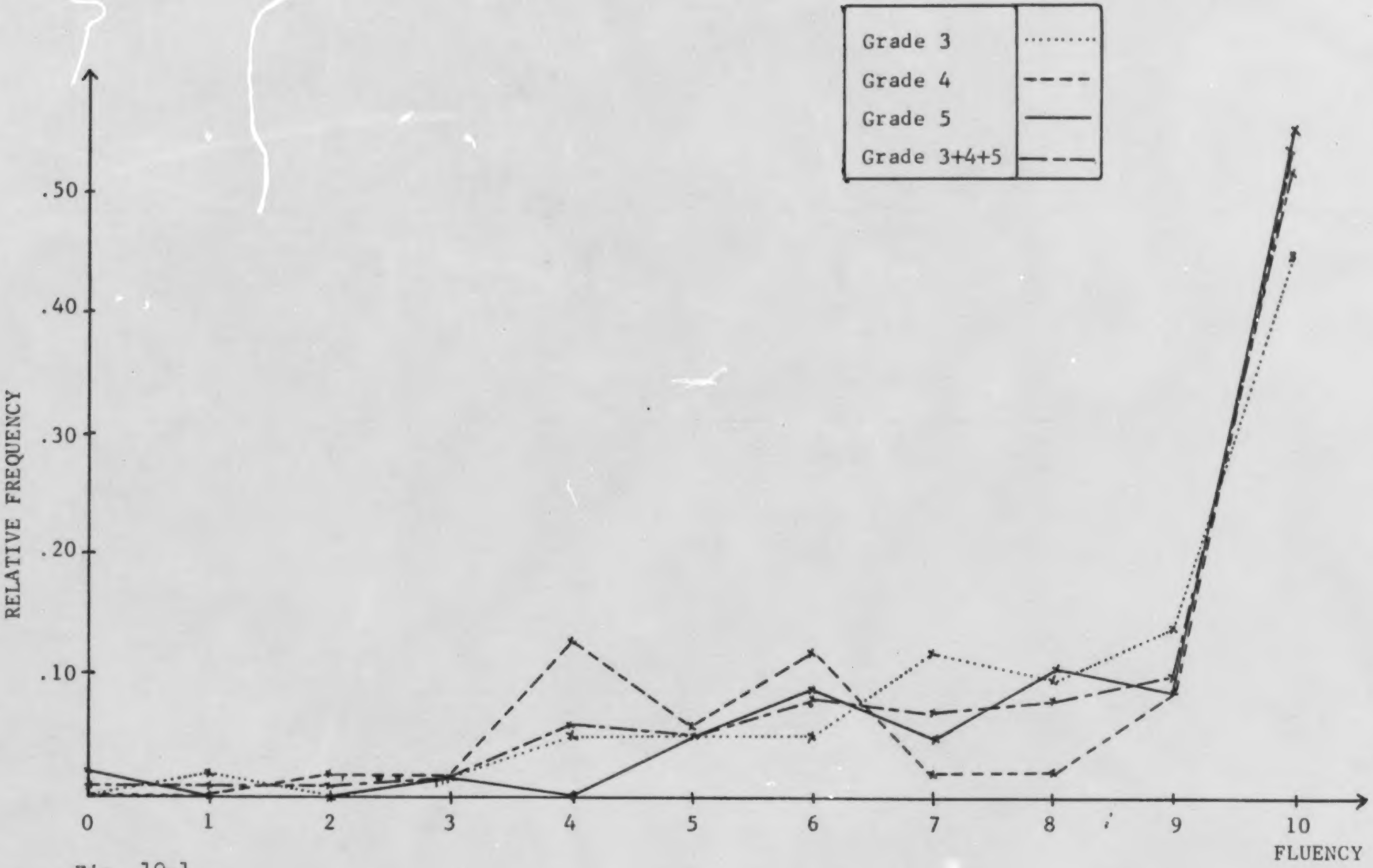


Fig. 19-1

The frequency distribution for different age groups in Torrance's Test of Creative thinking (figural part)

Activity 2: Fluency

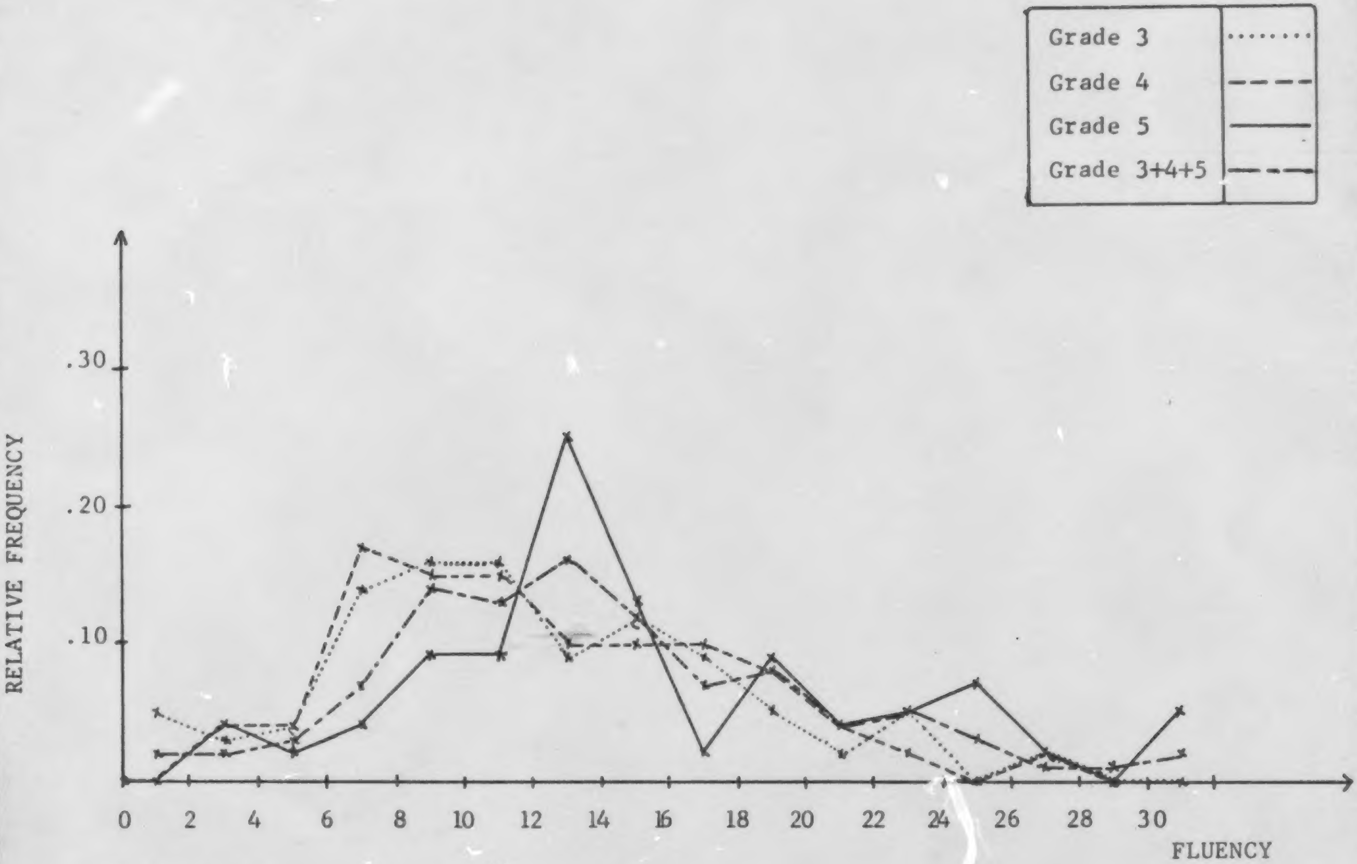


Fig. 19-2

The frequency distribution for different age groups in Torrance's Test of Creative thinking (figural part)

Activity 3: Fluency

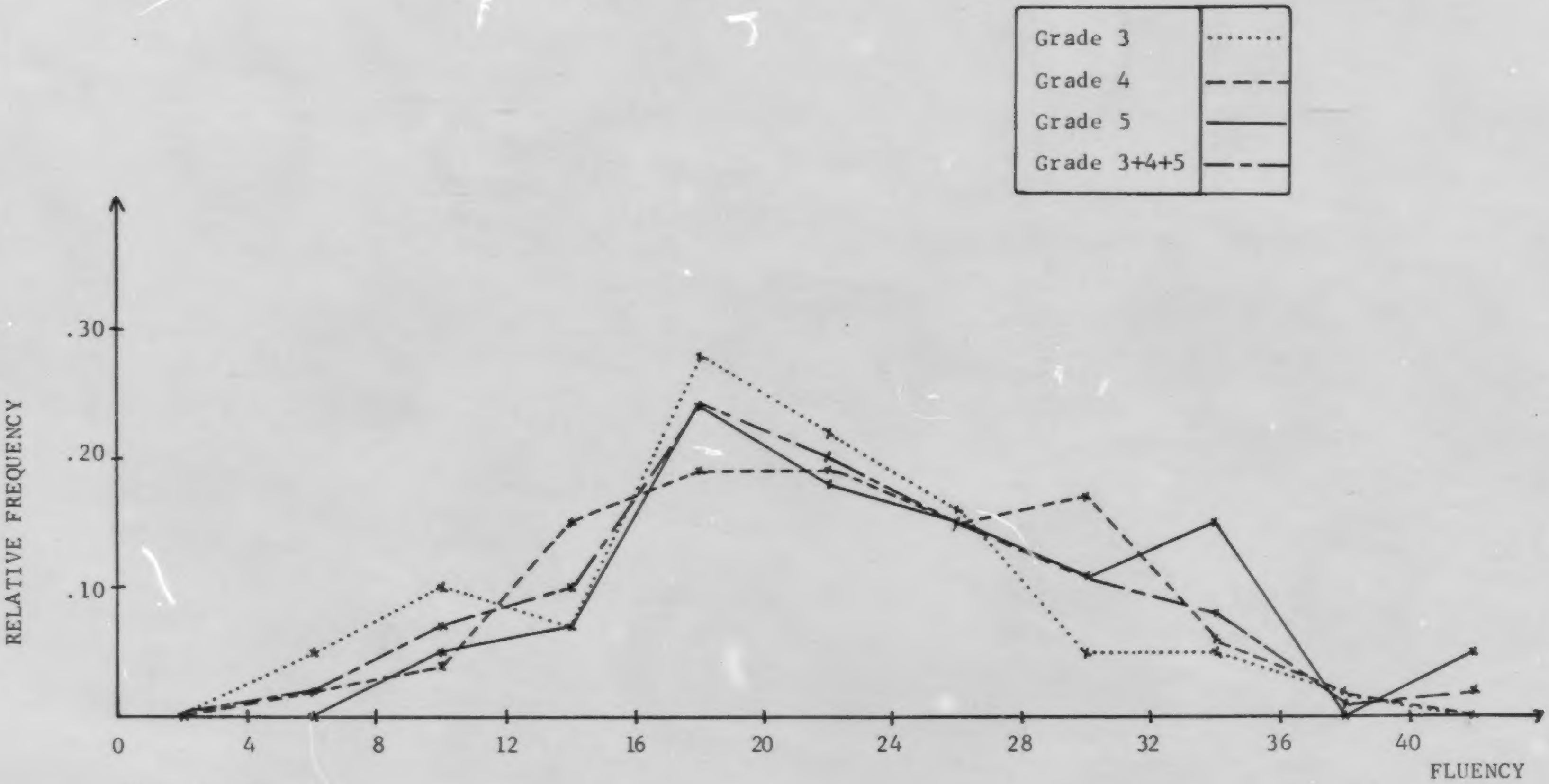


Fig. 19-3

The frequency distribution for different age groups in Torrance's Test of creative thinking (figural part)

Activity 1 + 2 + 3: Fluency

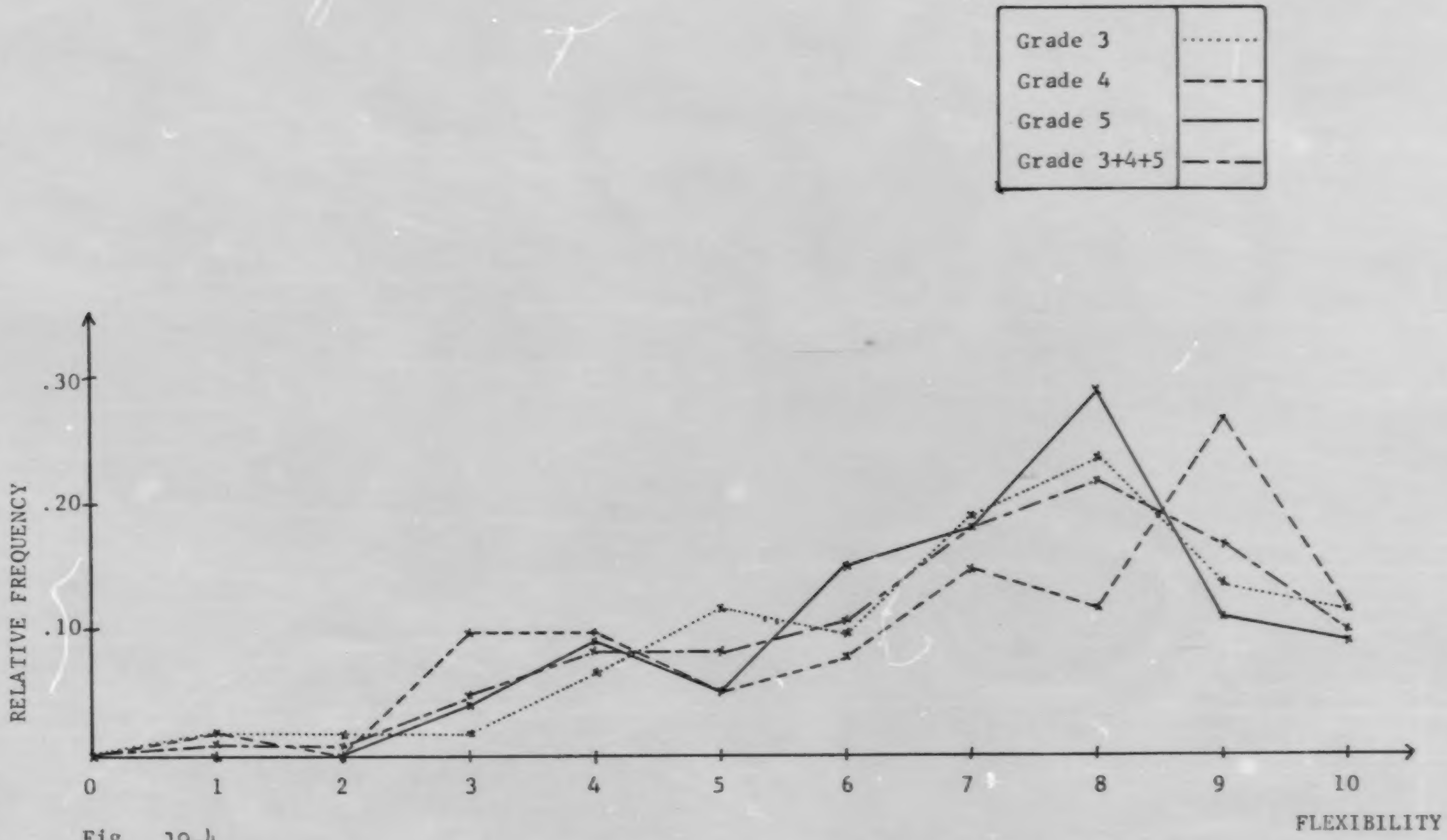


Fig. 19-4

The frequency distribution for different age groups in Torrance's Test of Creative thinking (figural part)

Activity 2: Flexibility

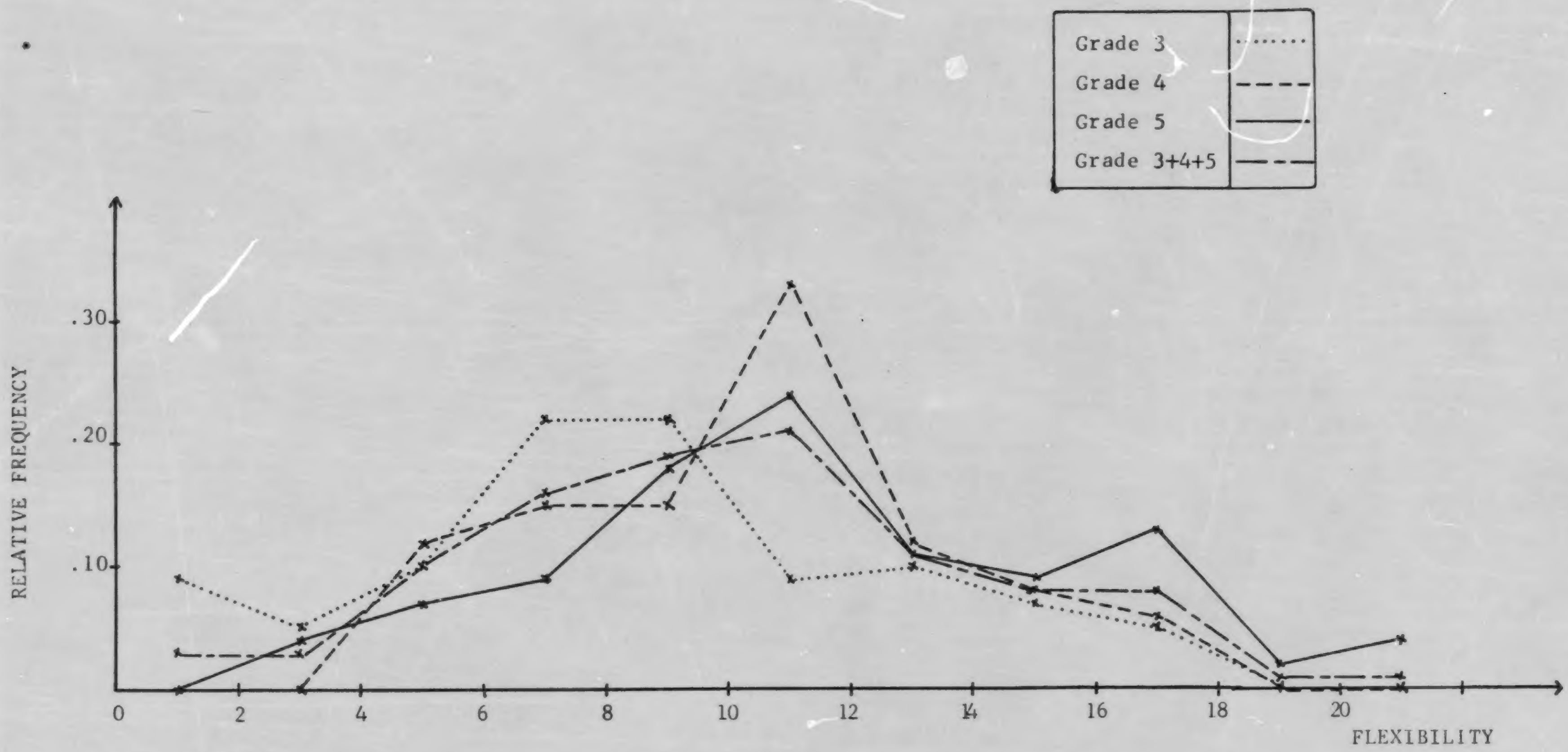


Fig. 19-5

The frequency distribution for different age groups in Torrance's Test of Creative thinking (figural part)

Activity 3: Flexibility

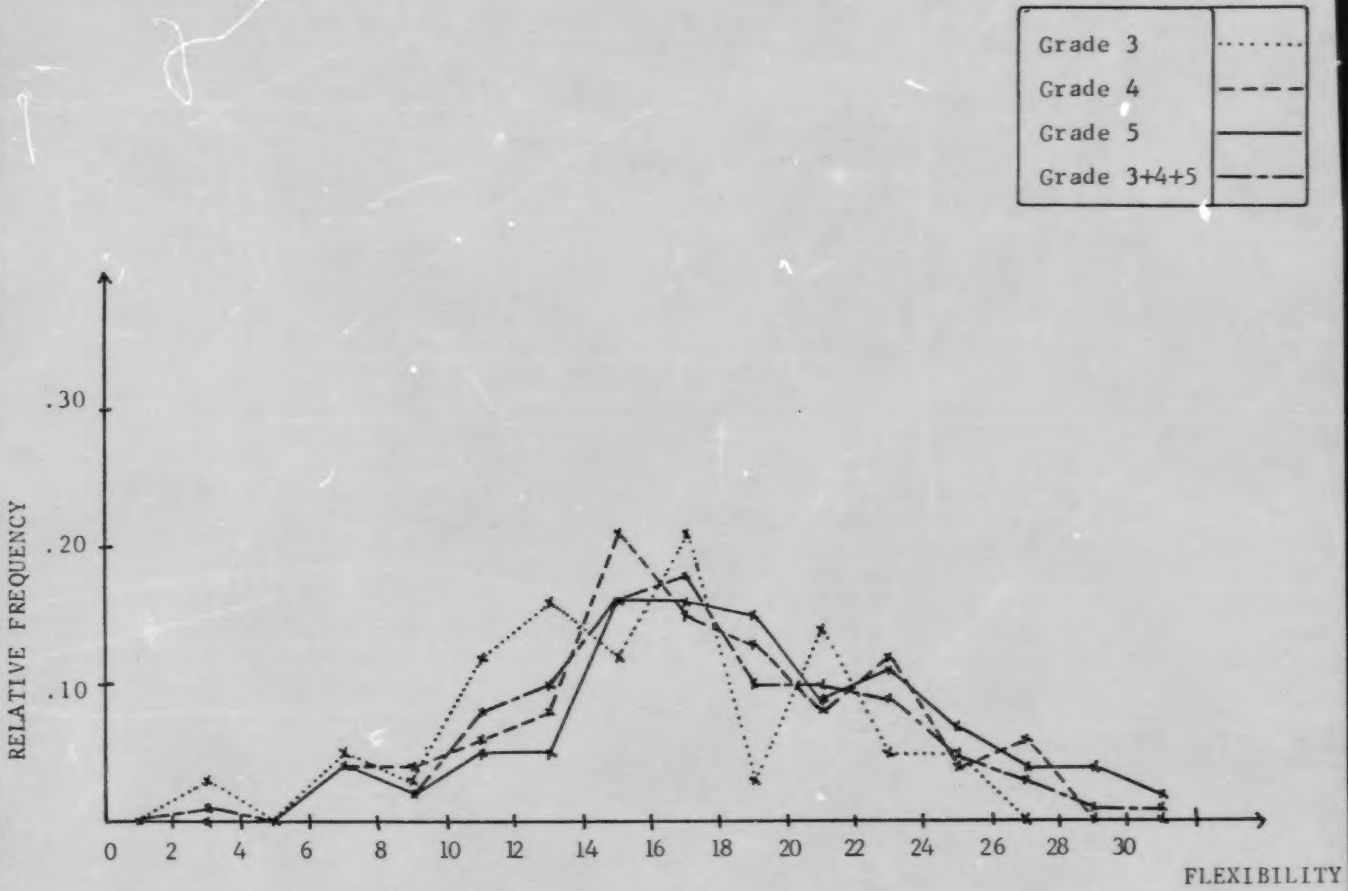


Fig. 19-6

The frequency distribution for different age groups in Torrance's Test of Creative thinking (figural part)

Activity 1 + 2 + 3: Flexibility

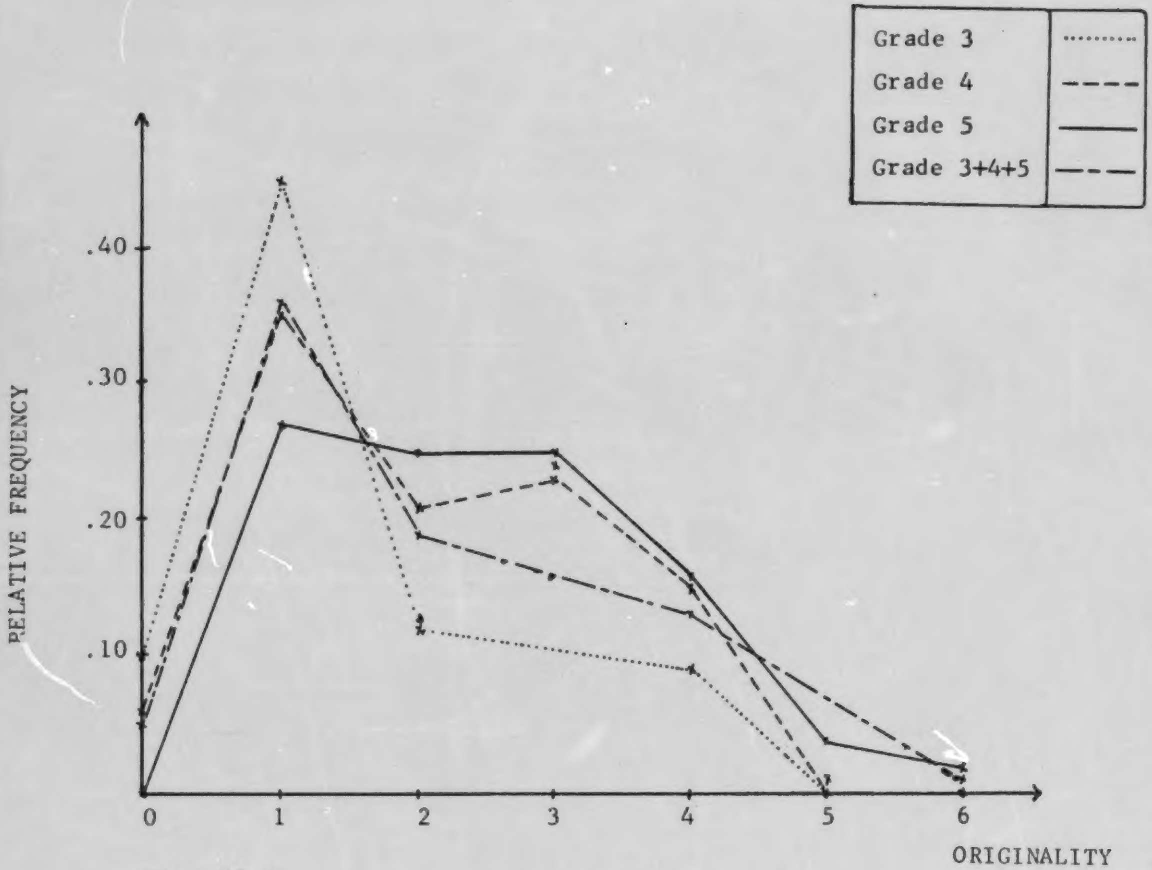


Fig. 19-7

The frequency distribution for different age groups in Torrance's Test of Creative thinking (figural part)

Activity 1: Originality $(O_i + O_{ii})$

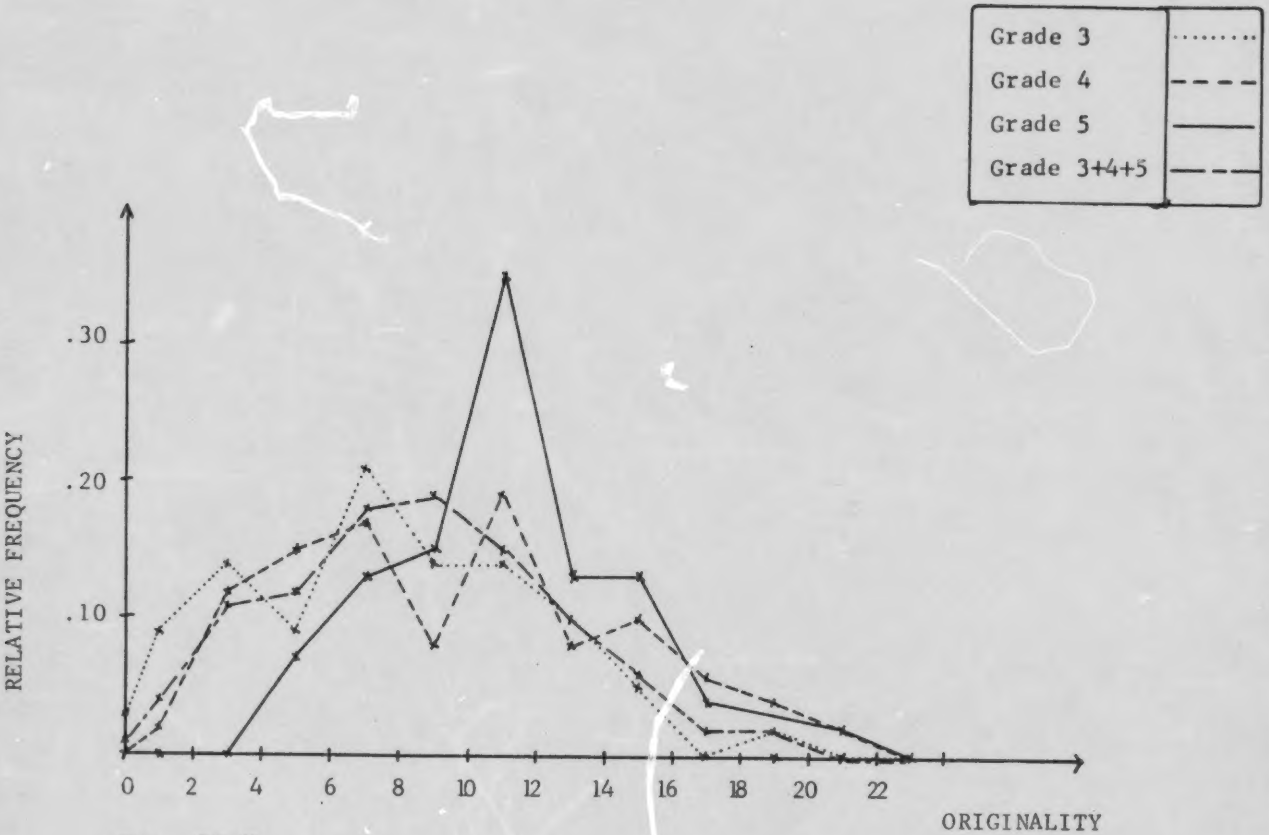


Fig. 19-8

The frequency distribution for different age groups in Torrance's Test of Creative thinking (figural part)

Activity 2: Originality ($O_f + O_{add}$)

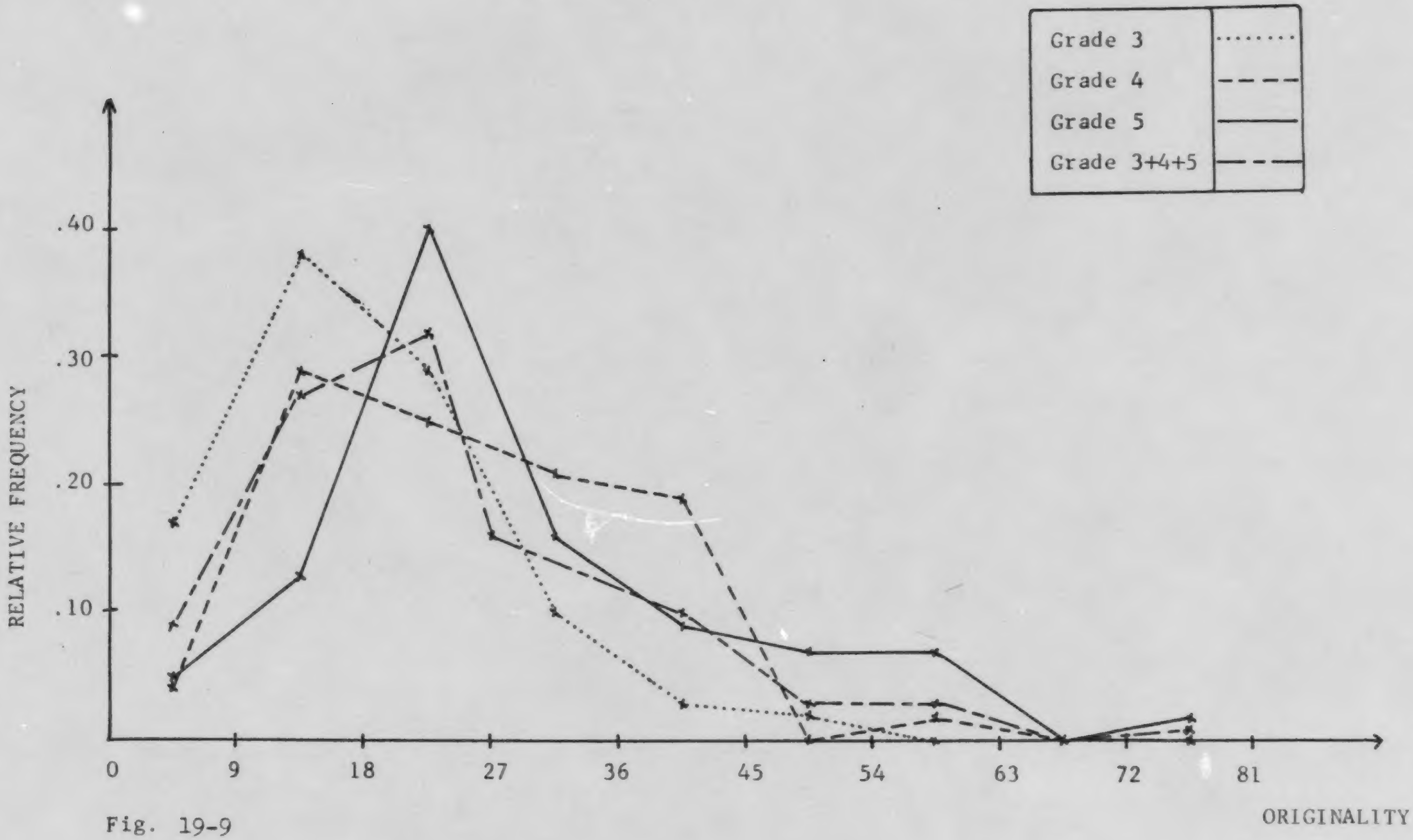


Fig. 19-9

The frequency distribution for different age groups in Torrance's Test of Creative thinking (figural part)

Activity 3: Originality ($O_f + O_{add}$)

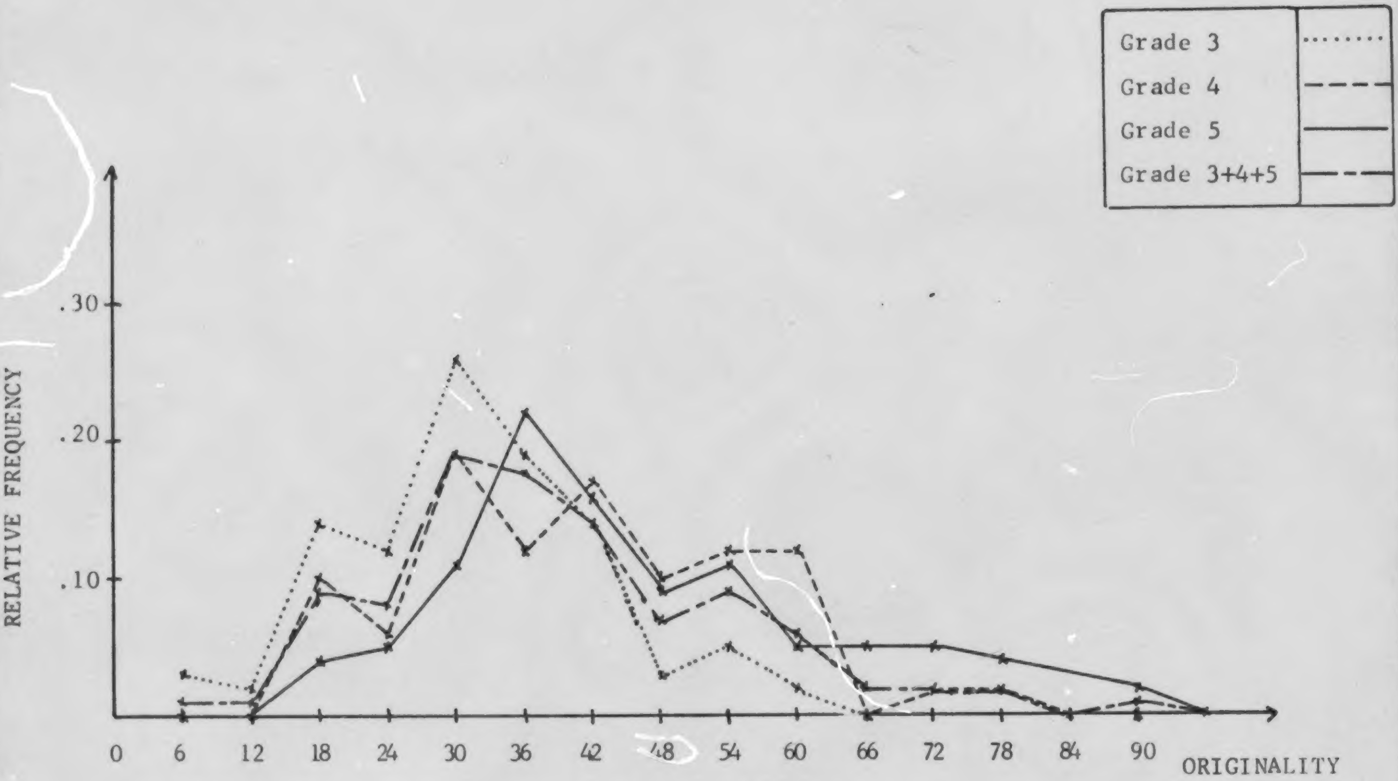


Fig. 19-10

The frequency distribution for different age groups in Torrance's Test of Creative thinking (figural part)

Activity 1 + 2 + 3: Originality $(O_t + O_{add})$

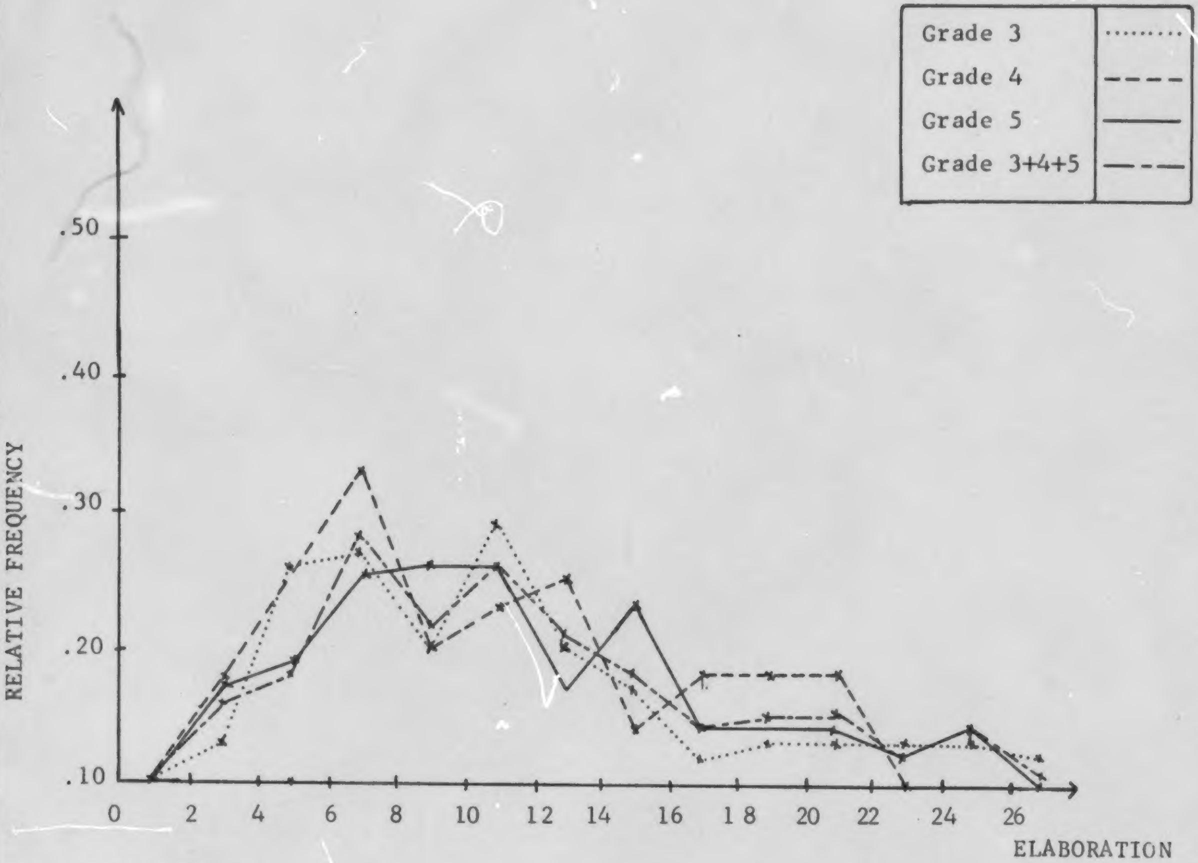


Fig. 19-11

The frequency distribution for different age groups in Torrance's Test of creative thinking (figural part)

Activity 1: Elaboration

Table 19-7. The mean difference between males and females in the Torrance Test of Creative Thinking, Form A, verbal part, product improvement.

Kind of score / Group	Number of students / Statistics	4		5		Σ[1]	
		male	female	male	female	male	female
		24	14	22	22	46	36
A (originality and fluency)	X̄	4.2	5.7	5.9	4.6	0.04	-0.04
	SD	3.06	3.41	4.47	2.28	3.85	2.88
	F	1.24		3.85*		1.79*	
	t	-1.32		1.18		0.08	
B (originality and flexibility)	X̄	2.0	3.4	2.9	3.0	-0.3	0.3
	SD	1.29	1.79	1.59	1.71	1.46	1.80
	F	1.93		1.16		1.51	
	t	-2.68*		-0.22		-1.91	
C (originality)	X̄	0.41	0.59	0.40	0.47	-0.09	0.09
	SD	0.17	0.14	0.14	0.15	0.16	0.15
	F	1.53		1.11		1.13	
	t	-2.57*		-0.71		-2.17*	
Creativity Kv	X̄	5.1	6.7	7.0	6.5	-0.2	0.2
	SD	2.67	3.29	3.54	2.33	3.14	2.81
	F	1.52		2.30*		1.25	
	t	-1.51		0.53		-0.63	
Fluency	X̄	12.6	10.9	14.3	11.1	1.1	-1.4
	SD	10.27	6.85	9.21	4.30	9.79	5.44
	F	2.25		4.59*		3.24*	
	t	0.59		1.43		1.34	
Flexibility	X̄	3.5	5.4	5.1	5.5	-0.4	0.6
	SD	1.91	2.44	2.15	2.44	2.05	2.49
	F	1.62		1.29		1.48	
	t	-2.39*		-0.52		-1.94	
Originality A	X̄	5.5	7.3	6.9	4.7	0.2	-0.2
	SD	6.44	4.20	4.85	3.56	5.81	3.98
	F	2.35		1.86		2.13	
	t	-1.06		1.74		0.38	
Originality B	X̄	0.5	0.7	0.5	0.4	-0.03	0.03
	SD	0.41	0.21	0.35	0.31	0.39	0.29
	F	3.76*		1.30		1.84*	
	t	-1.84		1.00		-0.36	

---

[1] See footnote 1 in Table 15-3

Table 19-8. The differences between the experimental and control classes in the raw scores of the French version of Torrance's Tests of Creative Thinking (Figural Part A).

			Activity 1				Activity 2				
			ORIG O_f	ORIG O_{add}	ORIG $O_f + O_{add}$	ELAB	FLU	FLEX	ORIG O_f	ORIG O_{add}	ORIG $O_f + O_{add}$
GRADE 3	EXPTL (N=20)	M	1.80	0.25	2.05	10.35	8.60	7.00	7.95	1.45	9.40
		SD	1.08	0.54	1.20	4.36	1.80	1.84	4.51	1.75	4.78
	CONTL (N=21)	M	1.19	0.38	1.57	9.33	8.24	7.19	6.76	2.00	8.76
		SD	0.96	0.72	1.14	4.98	1.97	1.89	3.96	2.16	4.42
		t	1.91*	-0.66	1.31	0.70	0.61	-0.33	0.89	-0.90	0.44
GRADE 4	EXPTL (N=20)	M	1.50	0.35	2.15	8.55	8.70	7.30	8.65	2.40	11.05
		SD	1.12	0.57	1.42	4.18	2.19	1.98	4.28	2.71	4.46
	CONTL (N=18)	M	1.67	0.56	2.22	13.11	7.06	6.39	7.11	1.94	9.06
		SD	0.94	0.76	1.13	5.21	2.84	2.75	3.60	2.07	4.42
		t	0.40	-0.93	-0.17	-2.96***	1.98*	1.16	1.20	0.59	1.38
GRADE 5	EXPTL (N=23)	M	1.96	1.22	3.17	11.87	5.61	7.13	7.30	4.00	11.30
		SD	1.08	0.98	1.20	4.56	2.50	1.75	2.90	2.48	2.63
	CONTL (N=21)	M	1.76	0.33	2.10	9.51	8.52	7.19	7.24	2.43	9.67
		SD	0.92	0.71	1.15	5.95	1.84	1.92	2.88	2.06	3.91
		t	0.64	3.45***	3.04***	1.46	0.13	-0.11	0.08	2.29**	1.62

*, ** and *** denote significance at the 10, 5 and 1 percent levels respectively.

Table 19-8 contd........

			Activity 3					Activities 1 + 2 + 3				
			FLU	FLEX	ORIG O_f	ORIG O_{add}	ORIG $O_f + O_{add}$	FLU	FLEX	ORIG O_f	ORIG O_{add}	ORIG $O_f + O_{add}$
GRADE 3	EXPTL	M	9.65	7.40	15.65	0.15	15.80	18.25	14.40	25.40	1.85	27.25
	(N=20)	SD	4.74	3.64	9.02	0.36	9.11	5.57	4.15	10.08	1.93	10.69
	CONTL	M	12.00	8.86	18.43	0.71	19.14	20.24	16.05	26.38	3.10	29.48
	(N=21)	SD	5.29	3.62	9.71	1.24	9.18	6.28	4.37	10.93	3.01	10.46
		t	-1.50	-1.29	-0.95	-2.00*	-1.17	-1.07	-1.24	-0.30	-1.59	-0.67
GRADE 4	EXPTL	M	15.30	10.05	24.20	0.75	24.95	24.00	17.35	34.65	3.50	38.15
	(N=20)	SD	4.99	2.97	11.10	0.94	10.35	6.75	4.29	14.51	3.54	13.86
	CONTL	M	12.56	10.33	24.78	0.61	25.39	19.61	16.72	33.56	3.11	36.67
	(N=18)	SD	5.66	3.92	13.20	1.16	13.04	7.88	6.22	16.67	2.96	16.84
		t	1.58	-0.25	-0.15	0.40	-0.11	1.83*	0.36	0.21	0.37	0.29
GRADE 5	EXPTL	M	17.87	12.43	35.26	1.17	36.43	26.48	19.57	44.52	6.39	50.91
	(N=23)	SD	6.33	4.13	15.72	1.56	15.44	7.98	4.92	17.17	4.24	15.49
	CONTL	M	13.10	9.76	24.14	0.71	24.86	21.62	16.95	33.14	3.48	36.62
	(N=21)	SD	6.57	4.32	13.85	1.12	13.72	7.59	5.85	14.68	2.77	14.18
		t	2.45**	2.09**	2.49**	1.09	2.63**	2.07**	1.59	2.37**	2.72***	3.20***

Table 19-9. The corresponding differences in the Torrance Tests, based on scores which were normalized separately for each grade.

			Activity 1		Activity 2		
			ORIG $O_f + O_{add}$	ELAB	FLU	FLEX	ORIG $O_f + O_{add}$
GRADE 3	EXPTL	M	103.05	100.41	100.87	99.71	101.42
	(N=20)	SD	13.61	10.80	11.12	13.32	15.88
	CONTL	M	98.46	96.96	99.04	101.62	99.57
	(N=21)	SD	12.90	12.44	12.06	14.03	14.42
		t	1.11	0.95	0.50	-0.45	0.39
GRADE 4	EXPTL	M	100.22	91.85	102.97	100.99	100.82
	(N=20)	SD	16.84	12.42	10.05	11.76	13.05
	CONTL	M	101.89	103.26	94.34	96.46	93.95
	(N=18)	SD	11.99	13.73	14.07	16.11	15.45
		t	-0.36	-2.67**	2.16**	0.98	1.47
GRADE 5	EXPTL	M	107.86	104.01	99.97	100.23	104.38
	(N=23)	SD	12.32	11.49	13.30	12.91	12.07
	CONTL	M	95.86	96.64	98.06	100.56	95.44
	(N=21)	SD	12.72	15.70	11.22	15.65	16.44
		t	3.17***	1.76*	0.52	-0.07	2.04**

Table 19-9 contd...........

			Activity 3			Activities 1 + 2 + 3			
			FLU	FLEX	ORIG $O_f + O_{add}$	FLU	FLEX	ORIG $O_f + O_{add}$	CREATIVITY
GRADE 3	EXPTL	M	96.30	97.90	97.29	98.58	98.81	100.59	99.32
	(N=20)	SD	12.26	12.73	12.94	9.59	9.29	9.09	8.18
	CONTL	M	101.79	102.44	102.04	100.42	102.03	100.02	100.82
	(N=21)	SD	13.67	12.64	13.66	10.38	10.11	8.49	8.41
		t	-1.36	-1.15	-1.14	-0.59	-1.06	0.20	-0.58
GRADE 4	EXPTL	M	104.69	101.00	100.01	103.83	101.00	100.35	101.73
	(N=20)	SD	13.36	13.47	15.38	11.00	10.54	10.72	9.83
	CONTL	M	96.19	101.63	100.59	95.27	99.04	98.81	97.71
	(N=18)	SD	16.40	17.37	17.44	14.35	15.68	12.76	13.63
		t	1.74*	-0.12	-0.11	2.05**	0.45	0.40	1.03
GRADE 5	EXPTL	M	106.50	105.46	107.41	103.23	102.85	106.55	104.21
	(N=23)	SD	12.49	13.50	12.81	11.72	10.42	6.87	8.65
	CONTL	M	95.11	96.09	95.77	96.56	98.32	95.69	98.86
	(N=21)	SD	15.44	15.34	14.14	11.04	14.36	8.51	10.20
		t	2.67**	2.14**	2.85***	1.89*	1.19	4.62***	2.56**

Table 19-10. The mean differences between experimental and control classes in the Torrance Test of Creative Thinking, Form A, verbal part, product improvement.

Kind of score / Group		4			5			Σ^1		
		Exp.	Cont.	Σ	Exp.	Cont.	Σ	Exp.	Cont.	Σ
	Number of students / Statistic	20	18	38	23	21	44	43	39	82
A (originality and fluency)	$\overline{X}$	4.16	5.44	4.77	5.23	5.34	5.28	-0.31	0.34	0.00
	SD	2.69	3.71	3.27	3.58	3.62	3.60	3.21	3.68	3.45
	F	1.90			1.02			1.31		
	t	-1.21			-0.10			-0.85		
B (originality and flexibility)	X	2.37	2.66	2.51	2.91	3.02	2.96	-0.09	0.10	0.00
	SD	1.27	1.99	1.66	1.65	1.66	1.65	1.49	1.82	1.66
	F	2.44*			1.01			1.49		
	t	-0.53			-0.23			-0.53		
C (originality)	$\overline{X}$	0.50	0.43	0.46	0.38	0.46	0.42	-0.01	0.01	0.00
	SD	0.14	0.19	0.17	0.13	0.15	0.15	0.14	0.18	0.16
	F	1.97			1.29			1.59		
	t	1.28			-1.97			-0.35		
Creativity	$\overline{X}$	5.08	6.32	5.67	7.04	6.47	6.77	-0.13	0.14	0.00
	SD	2.67	3.23	3.01	2.99	2.99	3.01	2.88	3.14	3.01
Kv	F	1.47			1.00			1.19		
	t	-1.29			0.62			-0.41		
Fluency	$\overline{X}$	9.00	15.28	11.97	13.61	11.76	12.73	-0.91	1.01	0.00
	SD	6.99	10.16	9.19	7.73	6.80	7.36	7.64	8.78	8.26
	F	2.11			1.29			1.32		
	t	-2.19*			0.84			-1.05		
Flexibility	$\overline{X}$	4.05	4.39	4.21	5.43	5.10	5.27	0.01	-0.01	0.00
	SD	1.94	2.63	2.30	2.36	2.24	2.31	2.18	2.44	2.30
	F	1.84			1.10			1.25		
	t	-0.45			0.49			0.05		
Originality 1	$\overline{X}$	5.05	7.33	6.13	5.52	6.10	5.80	-0.65	0.72	-0.00
	SD	3.23	7.51	5.79	4.88	3.78	4.40	-0.65	5.82	5.09
	F	5.39*			1.67			1.91*		
	t	-1.15			-0.44			-1.21		
Originality 2	$\overline{X}$	0.65	0.53	0.59	0.40	0.57	0.49	-0.02	0.02	0.00
	SD	0.38	0.33	0.37	0.30	0.34	0.34	0.35	0.35	0.35
	F	1.31			1.29			1.01		
	t	1.09			[illegible]			-0.43		

---

[1] See footnote 1 in Table 15-3

Table 19-11

Summary of the results of student creativity in the Sherbrooke Study.

ACTIVITIES	1				2					3					1 + 2 + 3						4							
FACTORS	ORIGINALITY			ELABORATION	FLUENCY	FLEXIBILITY	ORIGINALITY			FLUENCY	FLEXIBILITY	ORIGINALITY			FLUENCY	FLEXIBILITY	ORIGINALITY			CREATIVITY	FLUENCY	FLEXIBILITY	FLUENCY + ORIGINALITY A	FLEXIBILITY + ORIGINALITY B	ORIGINALITY			
GRADES	O_f	O_{add}	$O_f + O_{add}$				O_f	O_{add}	$O_f + O_{add}$			O_f	O_{add}	$O_f + O_{add}$			O_f	O_{add}	$O_f + O_{add}$						C	O_A	O_B	K_v
3	+												-								▨	▨	▨	▨	▨	▨	▨	▨
4				⊟	⊞					⊕					⊞						-							
5		+	⊞	⊕				+	⊕	⊞	⊞	+		⊞	⊞		+	+	⊞	⊕								
3 + 4 + 5	+		⊞		⊕				⊞	⊕					⊞				⊞	⊕								

+ Judged on raw scores the experimental group is significantly better

⊕ Judged on normalized scores the experimental group is significantly better

⊞ Judged on raw and normalized scores combined the experimental group is significantly better

- Judged on raw scores the control group is significantly better

⊟ Judged on raw scores and normalized scores combined the control group is significantly better.

Chapter 20

Perseveration and Sociometric Structure

20.1 Perseveration

Hypothesis 7: The Dienes method of mathematics learning results in less perseveration in the student than the traditional methods.

20.1.1 Theoretical background*

One speaks of perseveration when a response persists in an inadequate way from one situation to another. This functional inertia has been demonstrated in the most diverse forms of conduct. An example of "perceptive perseveration" is found in the delay in the appearance of grey on coloured disks (Wiersma, 1906; Cattell, 1946). Alternating writing tasks show "motor perseveration" (Cattell, 1946; Zazzo and Stambak, 1969). One may also speak of "mental and emotional perseveration", to name only the most generally recognized domains.

There has been much discussion of the problem whether there is a general characteristic of perseveration, or whether, on the contrary, the examples just cited correspond to independent realities. The earlier theories of perseveration — for example, the "secondary function" theory treating the ulterior effect of an exercise (G. Gross in "Die cerebrale Secundärfunction", 1902; see Spearman, 1927), or later theory emphasizing "a general inertia or mental slowness" (Spearman, 1927) - suppose the existence of a general factor P. When speaking of theories hypothesizing the existence of a general factor P, Cattell (1935) stresses that they consider a factor independent of the perseveration which can appear in intellectual activities.

---

* After Kónya, 1971.

He believes he can identify the mental perseveration studied by Muller and the associative perseveration analysed by Jung with the phenomenon of fluency. In the course of comparative analyses, mental perseveration (F) and motor perseveration (C) were found to be independent. Cattell (1946) includes the different manifestations of perseveration under the idea of "disposition-rigidity". The tests capable of measuring it include:

1. alternating tasks (performing two simple tasks alternately);
2. effort tasks (opposing an habitual activity and a new activity);
3. perception tasks (e.g. colour mixing).

He sums up the results of these tests by concluding that an individual having a high index of P cannot change his habits at will. According to the data of Cattell (1935) and Pinard (1932), based on the results of questionnaires and on observation by experienced observers, perseveration, which is seen as a functional characteristic, can be considered as a personality factor. Low values of the index (except extreme values) correspond to more balanced personalities. These people react more sensitively to stimuli from the outside world; they are less profound emotionally; and they change moods less easily. They have a more disciplined character, a more integrated personality, a more highly developed conscience. It also seems that perseveration in behaviour may occur as a consequence of anxiety.

We based our measure on Zazzo and Stambak's test of perseveration (1969), with which it has been demonstrated that the P index is not a function of age, since average results of adults and children are the same. This is not to say that an individual's P index cannot vary with time. On the contrary, the authors remark that the index of P can lessen under the influence of teaching and of education.

Our interest in linking tests of perseveration with the teaching of mathematics also developed from a remark by Cattell (1946) that students with a lower perseveration index seem to be attracted by mathematics, the technical sciences, and the natural sciences, while those with a higher perseveration index are drawn more by history, languages, literature and religion.

It would seem reasonable that the Dienes Programme because of its continually changing methods and materials should cause a lessening of the degree of perseveration.

20.1.2 Budapest study

METHOD

(Perseveration was examined in the Budapest and Sherbrooke studies, mainly by the same method. Here we describe the method in detail and in the next section only those aspects which were different in the Sherbrooke study.)

To measure the level of perseveration we used the Hungarian adaptation of the test of Zazzo and Stambak (1969). The test includes four graphmotor tasks (writing letters, numbers, sentences), and one verbal task (designating colors), which are all performed in both the ordinary way and in an unusual way. Perseveration is measured as a function of the slowing down accompanying the change to an unusual method (e.g. the diminution of the amount of written letters in 15 seconds, when first writing capitals and small letters, ABC, abc, and then the same letters in alternation, aA, bB, cC.) In view of the comparison which was made between this test and Cattell's (1935) P questionnaire, we can consider that the test indicated the presence of the following characteristics:

a. slowness of intellectual processes;

b. attachment to security, order, and habit (conformism);

c. emotional inhibition of activity, and exaggerated sensitiveness.

Comparison of the P index and behaviour characteristics shows that an average P index accompanies a certain ability to adapt. As, according to Zazzo and Stambak, children do not attain automatism in writing before 9 years of age (before which age graphomotor aptitude predominates in the measurement rather than perseveration), the lower age limit of our experimental population was grade 3. The upper limit was grade five, as the OPI method had not yet spread to higher classes. Because of the character of the test it was given individually and usually lasted about 15 minutes. We used the authors' instructions (which emphasize rapidity) as well as their method of calculating the P index.

TREATMENT OF DATA

The Mann-Whitney test was used to compare the results of the experimental and control classes. The Spearman rank correlation coefficient was used to correlate perseveration scores with the creativity results: fluency, flexibility, originality, and the total "creativity" ranking derived from the sum of the student ranks on the three components.

EXPECTATION

Smaller P indices in experimental classes than in control classes would indicate less rigidity of thinking, thus supporting hypothesis 7.

RESULTS

The average P indices are shown in Figure 20-1. Consistently less perseveration was found in the experimental classes than in the control classes although this difference became significant only in grade 5. The results support our hypothesis.

The correlations between the P index and the different parameters of creativity are shown in Table 20-1. The students' creativity is slightly connected with their perseveration, i.e., greater creativity is associated with smaller perseveration. This reaches the significance level in the grade 5 experimental class.

20.1.3 Sherbrooke study

METHOD

The method used here was the same as the one in the Budapest study, the French version of the perseveration test by Zazzo and Stambak (1969) being used. Because the testing was carried out in groups, the verbal subtest (designation of colors) was omitted, and the instructions to subtests were slightly altered from the original.

Subjects: Although the perseveration test was administered to all classes, the results for grade 3 were not considered, as more than half of the subjects did not meet the speed requirement set by the authors to ensure that the P index resulting from the test was in fact a measure of perseveration. In grades 4 and 5, results of all subjects were not usable for partly the same reasons, partly also because instructions were not followed well enough.

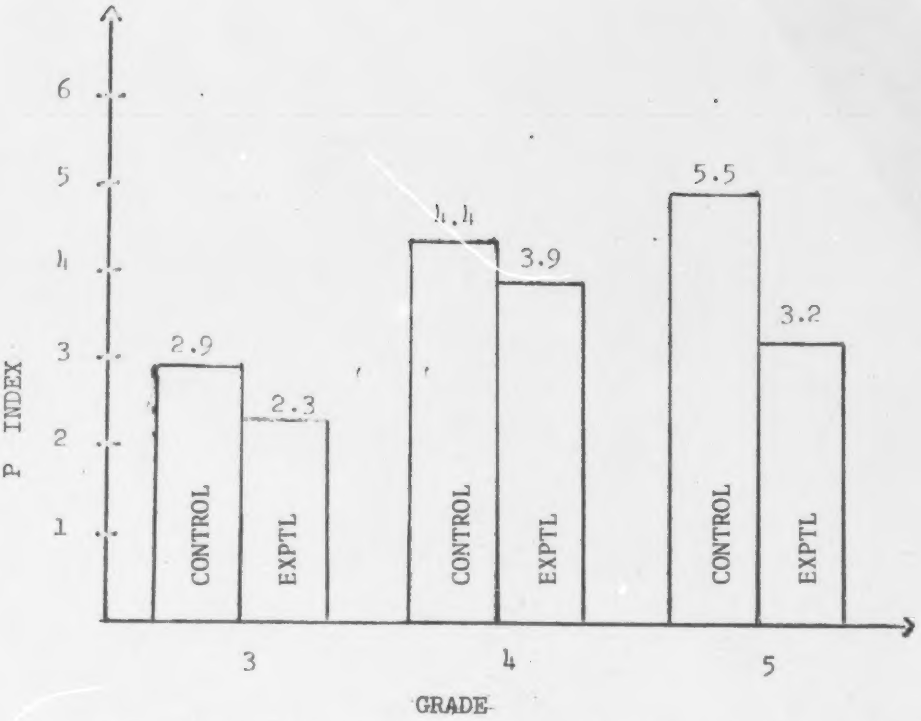


Fig. 20-1 The results of the examination of perseveration in the Budapest Study:

Table 20-1

The correlation coefficients between P index and the different parameters of creativity in the Budapest Study.

Classes		Fluency	Flexibility	Originality	Creativity	Number of subjects
3	Experimental	-0.235	-0.157	-0.147	-0.244	19
	Control	0.137	0.122	-0.147	not available	18
4	Experimental	-0.361	-0.171	-0.237	-0.275	18
	Control	0.036	0.042	-0.143	0.133	20
5	Experimental	-0.300	-0.089	-0.323	-0.561	13
	Control	-0.091	0.125	-0.179	-0.047	22

TREATMENT OF DATA

Raw scores for every child in each of the four subtests were calculated after the authors' instructions and then standardized for grades 4 and 5 together (comparison of these two grade levels showed that they were significantly different with $p < .05$ in two of the four subtests). The sum of the standardized subtest scores gave us the perseveration index P.

RESULTS

As Table 20-2 shows, there were no significant differences between the experimental and control classes. The tendency of students in the control group having a higher P index is to be seen in all three grades. Sex differences were checked and found to be not significant, and therefore had no influence on the results (Table 20-3). However, our hypothesis that students following the Dienes mathematics programme in Sherbrooke have a lower perseveration index has not as yet been confirmed by this study.

20.2 Sociometric structure

Hypothesis 8: The Dienes method of mathematics learning gives children greater opportunity to know each other during the educational process and within a more unified social structure.

20.2.1 Theoretical background

The OPI mathematics project* puts an emphasis not only on individualized learning, but also on the opportunity for group activity. Much of the activity in this arrangement is in various forms of group interaction, which may be

---

* The variable "social structure" was only examined in the Budapest Study.

Table 20-2

The mean differences in Zazzo's Perseveration Test between experimental and control classes in grades 4 and 5 and combined in the Sherbrooke Study.

Grades	Statistics Groupes	$\overline{X}$	SD	N	t
5	EXP	-0.08	1.82	18	-0.33
	CONTL	0.20	3.06	18	
4	EXP	-0.32	1.77	18	-0.88
	CONTL	0.21	1.74	16	
5 + 4	EXP	-0.20	1.80	36	-0.76
	CONTL	0.20	2.53	34	

Table 20-3

The mean differences in Zazzo's Perseveration Test between girls and boys in the Sherbrooke Study

SEX	BOYS			GIRLS				
Statistics Grades	$\bar{X}$	SD	N	$\bar{X}$	SD	N	t	F
5	+0.1	1.68	22	-0.3	1.57	22	0.64	1.14
4	-0.2	2.60	24	0.4	2.07	14	-0.83	1.57
Σ	-0.1	2.22	46	0.1	1.81	36	-0.24	1.51

propitious for the children knowing each other better than in the ordinary traditional classroom. Moreover, the project, de-emphasizing grades, rewards and punishments, stresses the inner motivation coming from the joy of recognition. So it seems to us a reasonable hypothesis that in this condition the children's judgement about each other is not distorted so much by envy, arrogance, inferiority, comtempt, etc.

20.2.2 Method

"Who are the three most clever pupils in your class?" -- a tailor - made test.

For measuring this aspect there exists a variety of methods, both traditional (e.g. Northway, 1952; Jennings, 1959; Reymond-Rivier, 1961) and newer (e.g. The Syracuse Scales of Social Relations; Gardner and Thompson, 1950), many of which were tried out also in classroom situations (Shoobs, 1947; MacKinney, 1948; Gronlund, 1959; Evans, 1962; Hoffman and Hoffman, 1966).

Because of the time limit we chose a very simple method. The children in grades 2 through 5 were asked to list (anonymously) the names of the three brightest children in the class, not necessarily those with the highest marks. The test was carried out (as Northway and Weld, 1957, suggest) individually in grades 2 and 3; and as a group test in grades 4 and 5. In each class the percentage of children naming the three most often named children was considered an inverse indication of the dispersion of their judgement.

20.2.3 Treatment of data

The results were graphed; no mathematical statistical methods were employed.

20.2.4 Expectation

Our expectation was that in the experimental class there would be a smaller dispersion of persons named by the children as the brightest amongst them than in the control class. According to our hypothesis a smaller dispersion would indicate:

i. better possibility for knowing each other during the educational process; and

ii. a more unified sociometric structure of the class.

20.2.5 Results

The results of this examination are visualized in Figure 20-1. As can be seen, the results unambigously show a smaller dispersion in the experimental classes of the persons named by the children as the brightest among them, thus supporting our hypothesis.

However, because of the lack of a complex sociometric examination the results of this test should be interpreted with great care. There is the possibility of other explanations of the observed phenomena; for example, greater pressure within the group could cause greater conformity and therefore greater congruity of the students' opinions. Everyday observation in the classroom, however, has led us to the interpretations mentioned above. In the future we should like to continue with more sophisticated investigation in this field.

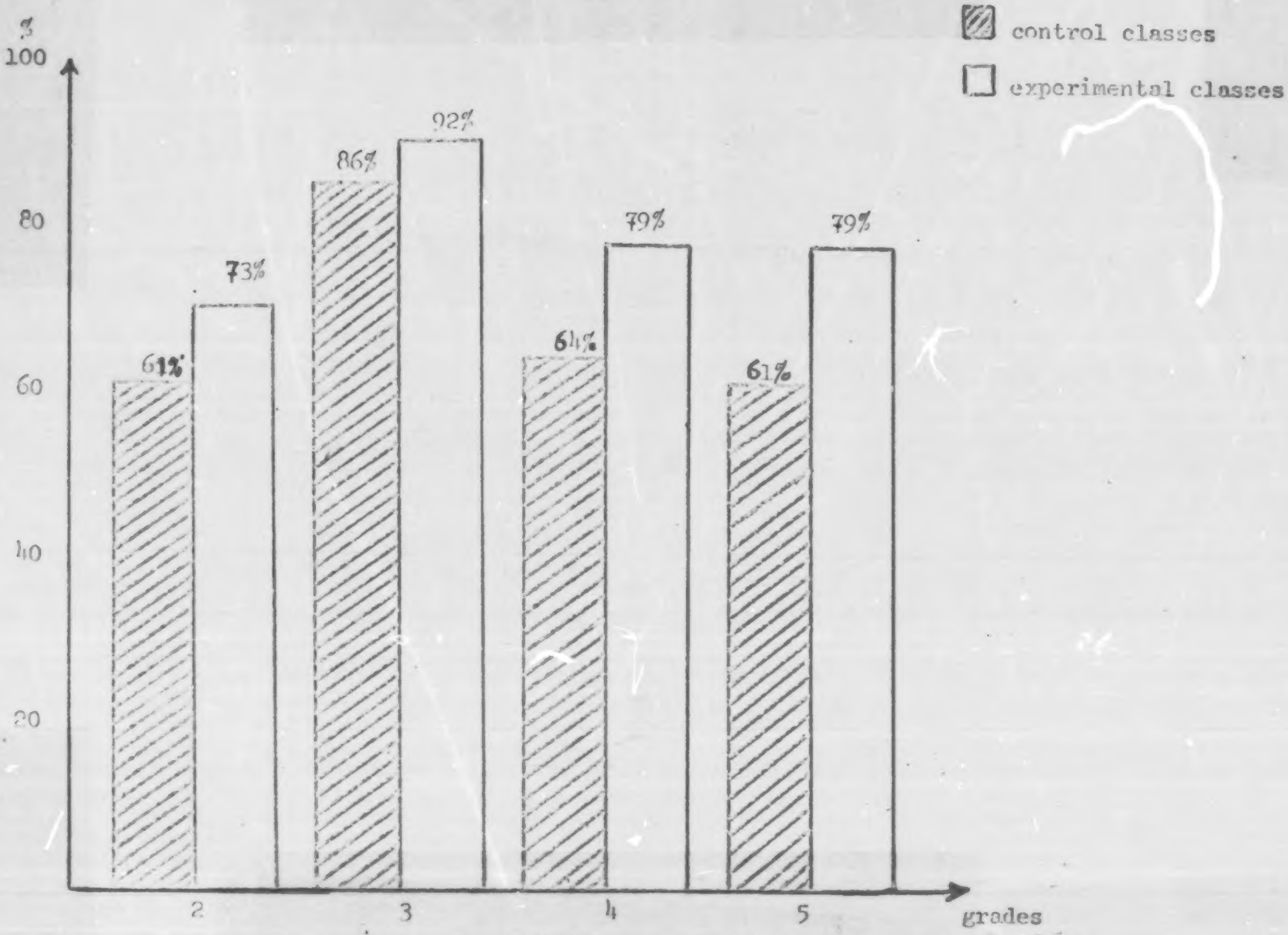


Fig. 20-2. Measurements of Hypothesis 8 in the Budapest Study. The percentage of children naming the three most often named children in their class. (An inverse indication of the dispersion of their judgement.)

Chapter 21

Attitude

Hypothesis 9: The students learning mathematics with the Dienes method have a more favourable attitude toward the mathematics lesson, and probably toward the whole of their school activities, than the students learning mathematics in traditional ways.

21.1 Theoretical background

The mathematics lessons in the classes using the mathematics teaching method were held in a freer, friendlier atmosphere. Our observation was that the students liked it and found it interesting. By measuring the students' attitude, we wanted to find proof of our hypothesis that the students' attitude toward the mathematics lesson is more positive in the experimental classes than in the control classes. We were interested as well in knowing if they tranfer this more positive attitude to the whole school.

We believe that attitude is one of the most important factors to be examined in an evaluation process. As Bloom (1969, p. 39) writes:

> If the curriculum, the evaluation methods and the standards that are used lead many students to regard themselves in a negative way or to develop a negative attitude toward school and learning, this is most unfortunate and should be investigated and, hopefully, corrected.

Most educators, at least when pressed, affirm the importance of enhancing curiosity and of implanting in the students massive and enduring positive affects toward learning and subject matter. Most of us would agree, therefore, that even when an instructional programme does not attempt to enhance positive attitudes directly, these variables should still be maintained if possible in the evaluation of the programme to guard against unintended decreases in interest or involvement (Messick, 1970).

On the other hand, it is quite difficult to find real differences in attitude between the experimental and control groups in educational studies, especially when young children are involved. For example, Hungermann (1967) contrasted ten classes who had used SMSG texts (see Volume I, Table 3-1) in grades 4, 5 and 5 with ten classes who had used a conventional text in these grades and found no differences in the attitude measures (while several differences were found in the achievement tests). Similarly, Worthen (1968), evaluating discovery and expository methods of teaching mathematical concepts to more than 500 fifth and sixth grade pupils, found no differences at a minimum acceptable level on measures of attitude using the semantic differential scale as well as a "statement attitude" scale (i.e. pupil's reaction to positive and negative statements about arithmetic). However, in another example, Wilson et al. (1972) were able to demonstrate a more favourable attitude towards "open plan" schools with 11 and 12 year old children.

The difficulty in the assessment of differences in young children's attitudes - evident though these differences may be through pure observation of their behaviour - is certainly to some extent due to the kinds of measurement the researcher has to refer to for economic and practical reasons.

Let us take, for example, the common paper and pencil group testing. We are not sure of the extent to which young children are able to abstract from their actual situation (mood, testing situation, immediately preceeding events) or make a fairly generalized judgement. Also, we are not sure to what extent they submit to response sets, especially to those of social desirability or to acquiescence, i.e. here giving responses in keeping with what their teachers or parents think or want them to think. Also, if we want to compare the attitude of different groups, it is not always sure that both groups have the same point of reference; some may like their school subjects because they like their teacher, or because they perform quite well (a relationship between attitude and performance in arithmetic has been shown by Davidson, 1950, in Biggs, 1962), while for others a positive attitude may be due to a real interest in the subjects. Because of these possibility interfering factors in attitude measurement with young children, each study on this matter has to be interpreted through the background of its own design and no general statement can be made yet as to the children's attitude toward modern mathematics. As for our studies, the same problems as described above may obscure the data, and because their complete elimination was beyond the frame of given possibilities, therefore, we must content ourselves with showing tendencies only.

21.2 Method

Osgood's semantic differential method

There are many different methods for measuring attitudes of pupils towards a school subject. For example, one of the first scales used to measure attitudes of pupils toward arithmetic was based on techniques invented by L.L. Thurstone (Thurstone and Chave, 1929; Dutton, 1954). Another popular method, the Likert-type attitude scale was also applied recently for this purpose (Dutton and Blum, 1968). Maertens (1968) used Osgood's semantic differential method for defining the effects of arithmetic homework upon the attitudes of third-grade pupils toward certain school-related structures. The necessity for economical testing and having young children as subjects made us choose the semantic differential as a method of attitude measurement.* It is an easy method to use for the experimenter as well as for the children. Also, since it refers more to feeling than to rational judgements and may be disguised as a word game, we assumed it to be the best method to counteract the problem of response styles mentioned above.

In general, the semantic differential is known and widely used as is a tool for the investigation of the connotative or affective meaning of concepts, where the term "concept" refers to all possible objects of judgement. In practical applications, the affective meaning of a concept is arrived at by having individuals rate it along graded scales (5, 7 or 9 steps), each having polar (opposite in meaning) adjectives as end points. The method was described in concordance with the theory of a conceptual space delimited by a set number of axes representing dimensions along which meaning could be said to vary. With respect to these axes, any given concept can be placed. Empirical evidence indicates that semantic space is defined by three axes: Evaluative, potency, and activity (Osgood et al, 1957). In recent studies by Osgood and his collaborators, this structure of affective meaning systems could be observed to be strikingly similar over different cultures; and based on scales of pan-culturally general factorial composition, the data for an atlas of the affective meanings of more than 600 concepts in 23 different language/culture groups could be gathered (Osgood, 1964; 1971). A valuable selection of studies using and evaluating this method is given in Snider and Osgood (1969).

---

* The validity of this method is attested by its high correlation with the traditional Thurstone, Likert and Guttman types of scales.

What is the relation between semantic space and attitude?

An attitude - as most authorities agree - is a disposition to respond to certain situations, persons, or objects in a consistent manner which has been learned and has become one's typical mode of response (Freeman, 1962, p. 596). According to Osgood, the most important points in the definitions of attitude are that it is "a learned, implicit process, which is potentially bipolar, varies in its intensity, and mediates evaluative behaviour" (Osgood et al., 1957, p. 190). From this point he continues to the conclusions that "attitude is the evaluative dimension of the total semantic space", and that attitude towards a concept is the projection of the point representing the meaning of the concept on to "the evaluative dimension of that space" (Osgood et al., 1957, p. 190).

For our purposes, we did not confine ourselves to the evaluation component, although to simplify the matter, we will refer to this testing procedure as an attitude measurement. Adding scales of dimensions other than of the evaluative one gave us a broader picture of what the children really felt. Also it reduced the possible effect of socially desirable responses, since the evaluative factor only is highly correlated with social desirability (Ford and Meisel, 1965).

The semantic differential is not so much a single measuring device as a general technique for obtaining and treating different kinds of measurement. In our three studies we applied this technique to different sets of scales and different kinds of measurement according to the prevailing conditions, as described in detail in every study.

21.3 Budapest study

21.3.1 Method

According to Osgood's theory, the most important steps of the attitude measurement procedure appropriate for group testing are:

1. Choosing a set of scales (pairs of adjectives) which have high loadings on the evaluative factor and negligible loadings on such other factors as potency, activity, etc.;

2. Displaying the adjectives "good" and "bad"; "interesting" and "boring", etc. at the two end points of a 7-step scale (or, as in our case, using children as subject, a 5-step scale);
3. Having subjects put a cross placing the concept (in our case the mathematics lesson and the school) somewhere between the two extremes;
4. Scoring the answer, giving 1 mark for the unfavourable poles of the evaluative scales and 7 marks (in 7-step scales) or 5 marks (in 5-step scales) for the favourable poles, regardless of the visual presentation of the scales to subjects which should be randomized in direction (in our case, we used the scoring system reversed, giving 1 mark for the answers in the favourable pole and 5 for the answers in the unfavourable pole);
5. Summing the evaluative ratings to obtain the attitude "score".

To study the students' attitude towards the mathematics lesson and the school we first chose scales which Osgood (1957, pp. 43, 53-61) analyzed by D-factorization (forced choice method) and square root factor analysis (thesaurus study) supplemented by some pairs of adjectives which we thought relevant to the concepts we were studying. The sets of scales are shown in Table* 21-1

---

* Notes on Table 21-1.

1. Our acknowledgements and thanks go to G.O. Nagy for his personal help in finding the Hungarian version of some of the scales originally used by Osgood et al.(1957).
2. The numbers in the left hand column show the position of the items in the Hungarian version, those in the right hand column the corresponding position in the English version.
3. The numbers in the right hand column also refer to two experiments of Osgood et al.(1957), the Thesaurus Study (pp.53-61 numbers without brackets only), and D-factorization, forced choice method (p 43). Items without a number did not appear in any of Osgood's studies, so that the factor to which they are assigned is merely hypothetical.

Table 21 - 1. Scales and Concepts used in Attitude Measurement in the Budapest Study.

Concept	Hungarian Order	Scales				English Order	Factor
		English		Hungarian			
School	1	good	bad	jó	rossz	1	evaluation
	6	serious	humorous	komoly	trefas	33	potency
	8	hard	soft	nehéz	könnyü	26	potency
	11	interesting	boring	érdekes	unalmas	58	receptivity
	16	fruitful	fruitless	eredményes	eredménytelen	--	evaluation (?)
	18	severe	lenient	szigorú	elnézö	28	potency
	20	noisy	quiet	zajos	csendes	--	activity
	25	sociable	unsociable	barátságos	barátságtalan	6	evaluation
	29	fair	unfair	igazságos	igazságtalan	(45)	evaluation
	31	excitable	calm	nyugtalan	nyugodt	38	activity
Mathematics Lesson	1	active	passive	szorgalmas	lusta	37	activity
	2	hard	soft	nehéz	könnyü	26	potency
	3	sociable	unsociable	barátságos	barátságtalan	6	evaluation
	5	tiring	refreshing	fárasztó	fihentető	--	potency(?)
	7	fair	unfair	igazságos	igazságtalan	(45)	evaluation
	9	fast	slow	gyors	lassu	41	activity
	12	unusual	usual	szokatlan	megszokott	53	novelty
	13	noisy	quiet	zajos	csendes	--	activity (?)
	14	good	bad	jó	rossz	1	evaluation
	15	severe	lenient	szigorú	elnézö	28	potency
	17	fruitful	fruitless	eredményes	eredménytelen	--	evaluation (?)
	19	excitable	calm	nyugtalan	nyugodt	38	activity
	21	successful	unsuccessful	sikeres	sikertelen	15	evaluation
	23	kind	cruel	jóindulatú	ellenséges	7	evaluation
	24	changeable	stable	változó	változatlan	44	novelty*
	26	serious	humorous	komoly	tréfás	33	potency
	28	interesting	boring	érdekes	unalmas	58	receptivity
	30	pleasant	unpleasant	kellemes	kellemetlen	(18)	evaluation
	33	loud	soft	hangos	halk	(16)	potency
	34	varied	unvaried	változatos	egyhangú	--	receptivity (?)
Play	10	collective	lonely	közös	magányos	--	----
	22	fast	slow	gyors	lassú	41	activity
	27	regular	irregular	szabályos	szabálytalan	--	----
	32	pleasant	unpleasant	kedves	durva	(18)	evaluation

---

* This scale appeared in Osgood's material in reverse order in the stability factor, with a loading of 0.19 . Its loading on the novelty factor was -0.11 .

As the Table shows and as Osgood recommended, we included a considerable number of scales supposed to represent factors other than the evaluative one, notably for the concept of play. These were not taken into consideration in scoring, but introduced in order to obscure the purpose of the measurement and to provide additional information on the meaning of the concepts of "school" and "mathematics lesson" for the students in the experimental and control classes.

Unlike Osgood we put only one scale on each page presented to the subjects, to avoid possible interference between scales. The pages were formed into a booklet which was easy for the students to handle.

Then, as there was no possibility of carrying out a factor analysis the following process of evaluation was used: First, all the scales concerning a concept were considered as though they were on one dimension. Summing up all 20 scales for the concept of mathematics lesson and separately the 10 scales for the concept of school, two "overall attitude scores" were obtained for each child. Then those scales were discarded which did not differentiate between the upper and lower quartile of students according to their overall attitude scores. Thus there were omitted the hard-soft, noisy-quiet, and excitable-calm scales, concerning the concept of school, and the hard-soft, tiring-refreshing and excitable-calm scales, concerning the mathematics lesson. (It is interesting to see that according to Osgood's studies none of these belongs to the evaluative factor.) The result of this process was a modified overall attitude score for each child, which was a basis for the evaluation of our hypothesis.

Secondly, we performed an elementary factorial evaluation. We carried out the same scale selection procedure as described above, but within each of Osgood's five supposed factors, and then left out from the evaluation those scales which did not follow the tendency in the factor. So, concerning the mathematics lesson, we had to leave out from this part of the evaluation the scales kind-cruel (evaluation), excitable-calm (activity), and the two pairs of adjectives supposed to measure the factor of novelty. Concerning the assessment of the school we could use only the scales supposed to measure evaluation and receptivity. This provided six modified factorial scores for each child, four for the mathematics lesson and two for the school, which complemented the overall score.

21.3.2 Treatment of data

The Mann-Whitney test was used to examine the differences in the modified overall attitude scores and in the modified factorial scores for the evaluation and activity factors (concerning the school and the mathematics lesson) and the potency factor (concerning only the mathematics lesson).

To give a feeling for the attitudes of students toward the mathematics lesson in the experimental and control classes, we have constructed a profile using the arithmetical mean of each class on each scale, including the ones which had been discarded for the modified scores discussed above (See Figure 21-1).

21.3.3 Expectation

Lower attitude scores in experimental classes than in control classes would show a more favourable attitude towards the mathematics lesson, or towards the school (depending on the part of the test considered), thus supporting hypothesis 9.

21.3.4 Results

The modified overall scores showed a consistently favourable attitude toward the mathematics lesson and toward the school, although the difference between the experimental and control classes reached the level of significance in both cases only in grade 4 (see Table 21-2).

The modified factorial scores were also considered. For the evaluative factor the experimental and control classes were significantly different in grades 3 and 4 when considering the mathematics lesson; the experimental classes leaned more towards "sociable, fair, good, fruitful, successful, kind, pleasant". Concerning the school the experimental grade 4 class leaned significantly more towards "good, fruitful, sociable, fair".

For the potency factor, in judging the mathematics lesson, the sum of ratings was larger in each experimental class (that is they leaned towards "refreshing, lenient, humorous, soft (not loud), soft (not hard)", though the difference did not reach the .05 level of significance in any grade.

In the activity factor, when judging the mathematics lesson, the experimental class was significantly closer to the side of "active, fast noisy, excitable" in grades 3 and 5.

Although these results support our hypothesis, without a longitudinal study it would be quite difficult to answer such questions as why significant differences were found mainly in grade 4.

Fig. 21-1 is an attempt at visualizing the results obtained in grades 3 through 5 concerning the mathematics lesson. It seems to give quite an

---

[*] Figs. 21-1 to 21-11 will be found at the end of this chapter.

Table 21-2. Results of the examination of the students' attitude towards the school and towards the mathematics lesson in the Budapest Study.

Classes	Evaluation		Potency	Activity	Modified overall scores	
	Maths Lesson	School	Maths Lesson	Maths Lesson	Maths Lesson	School
3	+	0	0	+	0	0
4	+	+	0	0	+	+
5	0	0	0	+	0	0

Key: + significantly better result in the control class
0 no significant differences between experimental and control classes.

impressive demonstration of the difference between the experimental and traditional mathematics lessons. (The experimental one is more active, soft, sociable, refreshing, fair, fast, unusual, noisy, good, lenient fruitful, excitable, successful, kind, changeable, humorous, interesting, pleasant, and loud; the control one, on the other hand, is more passive, hard, unsociable, tiring, unfair, slow, usual, quiet, bad, severe, fruitless, calm, unsuccessful, cruel, stable, serious, boring, unpleasant, and soft.) But because there is no evidence that the differences along the scales are significant, any conclusion derived from this presentation should be handled with great care.

21.4 Sherbrooke study

21.4.1 Method

The instrument used in the Sherbrooke study was in all but one aspect the same as the one used in the Budapest study, the only changes being in the amount and the nature of the scales. Thirty-four instead of twenty-four scales were chosen to have the concept "mathematics lesson" judged, but only six scales (instead of 10) for the concept "school", 4 scales (in both studies) for the concept "play". The concepts "school" and "play" are here of minor interest. The reason for having them included is to obscure somewhat the purpose of measurement, to loosen up the child's situation of continuous judgements and to provide him with a kind of frame of reference. Table 21-3 shows the sets of scales in French and in English for the three concepts, their order of presentation in the test booklet as well as their factorial origin from Osgood's thesaurus study (1957).

Table 21-3

The 34 scales as used in the measurement of attitude towards

The Mathematics Lesson ("La leçon de mathématique"),

School ("L'école") and Play ("Les jeux") in the Sherbrooke Study

No.	Order in the test booklet	French		English		Factorial Origin
		The Mathematics Lesson				
1*	1	difficile	facile	difficult	easy	potency
2	2	on y parle beaucoup	on n'y parle pas	talkative	quiet	
3*	4	pas valable	valable	valueless	valuable	
4	5	on y fait beaucoup de choses /	on ne fait rien	we do alot	we don't do anything	activity
5*	6	fatigante	reposante	tiring	restful	
6	8	juste	injuste	fair	unfair	evaluation
7*	10	pas nécessaire	nécessaire	unnecessary	necessary	
8	11	rapide	lente	fast	slow	activ./pote
9*	13	toujours pareille	différente	always the same	different	novelty
10	14	bruyante	calme	noisy	calm	
11*	15	sans aucun sens	a du bon sens	meaningless	meaningful	evaluation
12	16	bonne	mauvaise	good	bad	evaluation
13	17	sévère	pas sévère	severe	lenient	potency
14	18	utile	inutile	useful	useless	
15*	20	pas importante	importante	unimportant	important	evaluation
16	21	agitée	tranquille	agitated	calm	activity
17	22	un succès	un désastre	successful	unsuccessful	evaluatio
18	24	humaine	inhumaine	kind	cruel	evaluation
19	25	simple	compliquée	simple	complex	activity
20*	26	ne change jamais	changeante	stable	changeable	stability
21*	28	sérieuse	gaie	serious	humorous	potency
22	30	pratique	sans valeur	useful	useless	
23	31	agréable	désagréable	pleasant	unpleasant	evaluation
24	32	intéressante	ennuyante	interesting	boring	receptivity
25*	35	imprudente	prudente	rash	cautious	stability
26	36	bruyante	silencieuse	loud	quiet	potency
27	37	variée	pas variée	varied	unvaried	
28*	38	ancienne	nouvelle	old	new	novelty
29	39	amicale	inamical	sociable	unsociable	evaluation
30	40	entrainante	endormante	stimulating	boring	
31*	41	vague	précise	vague	precise	
32	42	heureuse	malheureuse	happy	unhappy	
33*	43	longue	courte	long	short	
34*	44	plate	amusante	dull	amusing	—
		School ("L'école")				
1	3	bonne	mauvaise	good	bad	
2	7	sérieuse	gaie	serious	humorous	potency
3	9	difficile	facile	difficult	easy	
4	19	sévère	pas sévère	severe	lenient	potency
5	27	amicale	inamicale	sociable	unsociable	evaluation
6	33	juste	injuste	fair	unfair	evaluation
		Play ("Les jeux")				
1	12	en groupes	seul	in group	alone	
2	23	lents	rapides	slow	fast	activity
3	29	inhabituels	ordinaires	usual	unusual	
4	34	agréables	desagréables	pleasant	unpleasant	evaluation

* These scales were reversed for data analysis.

The scales which do not have this factorial indication are those chosen by us. In choosing these scales, consideration was given to relevancy of the concept being tested and the likelihood of its being understood by young children.

Subjects: The test was administered in all grades and results were obtained from 165 subjects. Comparisons were based on 123 subjects.

21.4.2 Treatment of Data

Because, for the Sherbrooke study, computer facilities were provided, the data were subjected to a factor analysis. A reorganization of the data by means of a factorial analysis is sensible insofar as the "evaluation-potency-activity structure" apparently can be expected to be obtained only when a large and heterogeneous concept sample is employed, but not when the sample is homogeneous (for the semantic differential in relation to multivariate research, see Miron and Osgood, 1966): In this study, only the concept "mathematics lesson" is considered and several scales which are not derived from the three factors were included. Therefore, the factors which were extracted in our case would reflect those dimensions (out of the provided set of scales) which were relevant for our group of children. The factor analysis was carried out in the three grades separately because somewhat different dimensions are to be expected in different age groups. One has to consider different levels of understanding and familiarity, also changes in the relevance of the scales for different ages.

Before factor analysis the scale values obtained from each individual were altered: Instead of running from 1 to 5 they became -2, -1, 0, 1, 2; some scales had to be reversed so that ratings on the favoured side (valuable, useful, etc.) received negative scores, while the "bad" sides (valueless, useless, etc.) received positive ones. In some cases we decided on the favoured side in accordance with our theoretical objectives: For example, "loud" on the favoured side received a low score and "quiet" on the other side a high one. These modified scores were then factor-analysed for every grade, with the result that we were able to obtain factor scores for every student instead of just a simple rating-score on every scale.

A factor score is a combined score of a student's ratings on the scales, weighted by the corresponding loadings of those scales on the factor. By taking factor scores we gave more importance to the dimensions of judgement of our group of subjects, as revealed by the factor analysis. In order to compare experimental and control groups we

a) made a profile of the ratings (Fig. 21-2, 21-3, 21-4), and

b) computed mean differences on the factor scores of 7 factors as well as on their sum.

A low factor score of a student indicates a favourable attitude in relation to the tables (21-4, 21-5, 21-6) showing the factor compositions; it means that the rating of "la leçon de mathématique" inclines towards the three adjectives given there.

21.4.3 Expectation

The students of the experimental classes should in general have lower factor scores than those of the control classes, indicating a more favourable attitude towards the mathematics lesson.

21.4.4 Results

By the applied factor analysis 7 factors were extracted and rotated. Tables* 21-4 to 21-6 show the 7 factors separately for each grade. Each factor is characterized by the scales which have the highest loadings (above .30) on it, ordered after their loading. Looking at the factors in the different grades it might be seen that though some factors are quite similar (see e.g. a simplicity factor in grade 3 as factor IV and in grade 5 as Factor V) they are not the same. Only the first factor, which may be labelled as an "interest - importance - happiness" factor can be said to be the same through all three grades. Considerations of space preclude the further description or interpretation of the factors. Let us only mention an effect which we caused by our assumption that "noisy" "loud" "agitated" are favourable attributes of a math lesson: In most of the factorial compositions (see factors II in grade 5, III and IV in grade 3) these traits are in the same line of judgement by our tested children in Sherbrooke as less favourable traits such as "long", "slow", "unnecessary". This effect may blur results when one makes comparisons on the sum of factor scores. Table 21-7 shows the differences on the factor scores between experimental and control groups, table 21-8 those between girls and boys.

As one can see, we found significant sex differences only in some cases. In grade 5 the boys found the mathematics lesson more favourable along Factor 5 (simple, fast, fair, stimulating, interesting, good, short, easy), Factor 7 (sociable, humorous, pleasant, kind, easy, etc.), and in the sum of all the seven factors. On the other hand, in grade 4 girls found the mathematics lesson happier, more pleasant, more interesting, better, more meaningful, etc., (that is more favourable along Factor 1) than did the boys.

---

* Tables 21-4 to 21-8 will be found at the end of this chapter.

As for the experimental-control differences the results are quite controversial. In grade 5 on Factor 1 (important, different, interesting, necessary, etc.), on Factor 3 (meaningful, good valuable, etc.) and also on the sum of the Factors the control class gave more favourable answers, while in grade 3 (interesting, stimulating, sociable, pleasant, etc.) and in grade 4 (happy, pleasant, interesting, good, etc.) on Factor 1 the attitude was more favourable towards the mathematics lesson in the experimental classes.

While it is difficult to generalize these results they could serve certainly as useful feedback - especially to the grade 5 experimental teacher.

21.5 New York Study

The objective of this study was to ascertain the attitude of children at the Fleming School, New York, towards:

(i) their school,

(ii) their teachers, and

(iii) their mathematics lessons.

21.5.1 Method

Osgood's method of the semantic differential again was used in this study in order to assess the children's attitude towards different school-related concepts. The method is described from a general point of view in section 21.2 and in its practical application under 21.3.1. By it the concept's affective meaning for the testee rather than the latter's cognitive judgement is assessed.

Out of our experiences with this method in the Budapest Study and later in the Sherbrooke Study we developed a new and improved set of scales . These scales were chosen according to two criteria:

(i) their natural relation to the 3 dimensions found empirically by Osgood and his collaborators (Osgood 1964, 1971) to make up the structure of the human affective meaning system (see 21.2), and

(ii) their frequency and diversity of usage by 9 year old children on different nouns (Di Vesta, in Snider and Osgood 1969).

The number of concepts was enlarged beyond those suggested by the overall objective to become:

(i) SCHOOL,

(ii) MATHEMATICS LESSON,

(iii) TEACHER,

(iv) MY BEST FRIEND, and

(v) PLAY.

These concepts were then presented to the children by means of a 7-step scale, which was found to be suitable for the age group we were testing. Owing to lack of a proper control group (in the Sherbrooke Study a different set of scales and a 5-step scale had been used) our analysis of the data could not exploit the full potential of the instrument. The concepts together with the scales on which they had to be rated were presented in the form of a small test booklet entitled

"What does it mean to you?"

The instructions emphasized the fact that this was a word game. No identification was asked from the children in order to allow them to express their attitudes freely.

21.5.2 Treatment of data

The obtained data were averaged over groups, i.e. the ratings of each concept on each scale were averaged once over the class as a group, once over the grade as a group, and once over both grades. In order to illustrate the children's affective attitudes and to give a basis of comparison profiles for each rated concept were constructed.

21.5.3 Results

The figures 21-5, 21-6 and 21-7 show the rating profiles for the concepts "Teacher" and "My best friend" in grade 4, grade 5 and grade 4 and 5 together.

Both ratings were put into one figure in order to show in which aspects they converged the most and in which the least. We wish to draw attention to the fact that the profiles of "Teacher" and "My best friend" are somewhat closer in grade 5 than they are in grade 4; this is especially true for the scales good-bad, wonderful-terrible, unfriendly-friendly. It is worthwhile to notice also that on the dimension of potency which is represented by the scales big-little, powerful-powerless, deep-shallow and weak-strong there is no difference between the ratings of "My best friend" and "teacher", though it might be expected from a cognitive physical point of view; this fact might indicate that "the teacher" (the children were asked to think of all teachers of the school simultaneously when rating this concept) is more a person of equal rank than a dominating authority for the Fleming school children.

Figures 21-8 to 21-10 show the profiles of the concept "school" "Math lesson" in comparison to "Play" for the same groups as above. The profiles of grades 4 and 5 together show that the concept "school" is rated somewhat nearer to "Play" than "Math lesson" on the evaluative dimension (scales: good, nice, sweet) and the activity dimension (scales: fast, alive). For the potency dimension all three concepts are quite near to each other; they are all considered as somewhat strong, big and powerful. Worthwhile to notice here is that "School" is seen as more helpful than "Play". Differences between the two grades consist of the higher extremes with which the concept "Play" is rated by grade 5 children, a fact which also generates greater distances from the other concepts on some of the scales.

Figure 21-11 displays the affective meaning which the children of each class have towards the concept "School". Particular observations and interpretations may be left to the reader.

The "yellow" classes seem to have a slightly more favourable attitude towards "School" than the others!

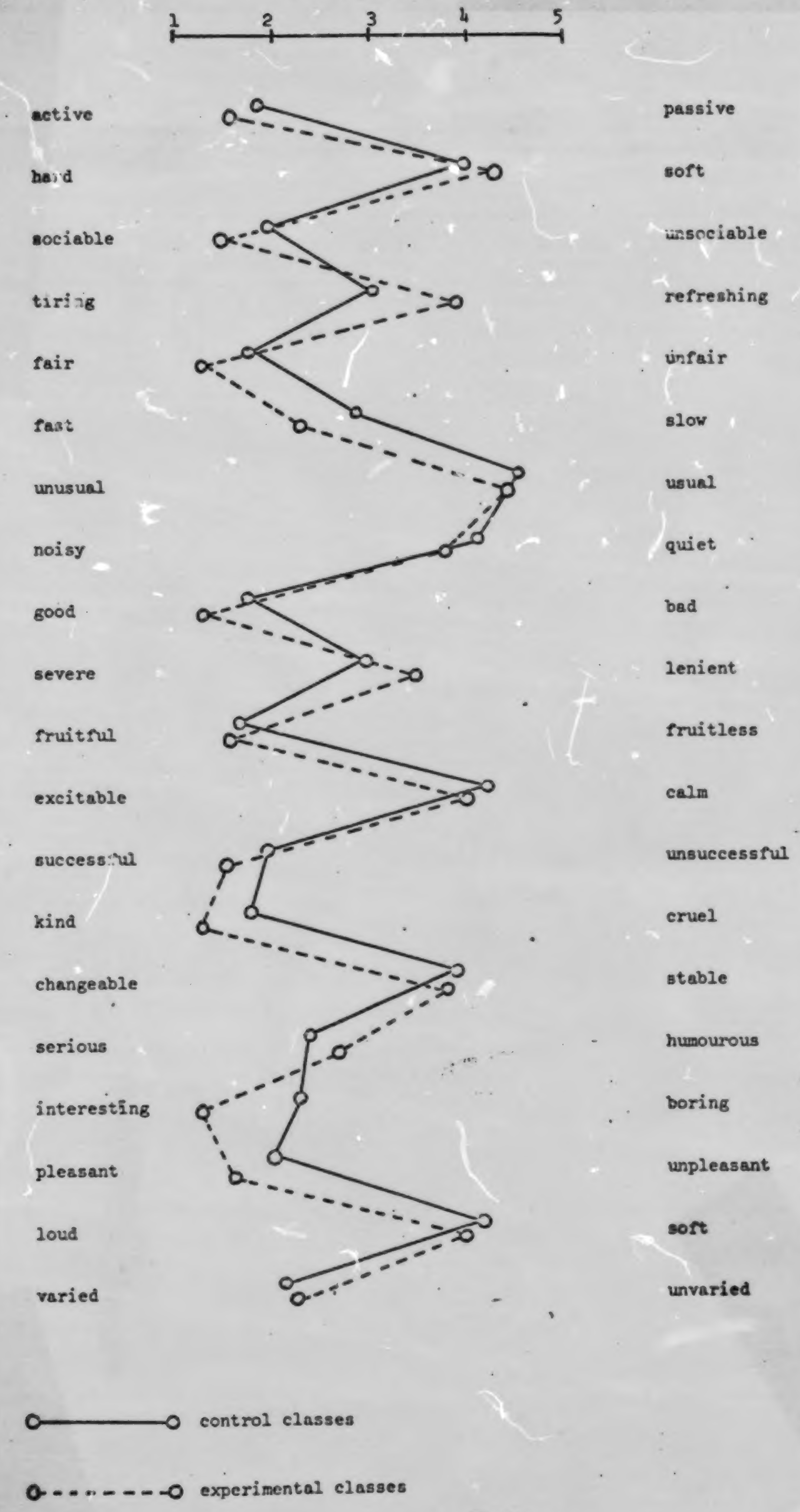


Fig. 21-1

The profiles of the experimental and control classes, rating the concept "mathematics lesson" against 20 scales in the Budapest Study.

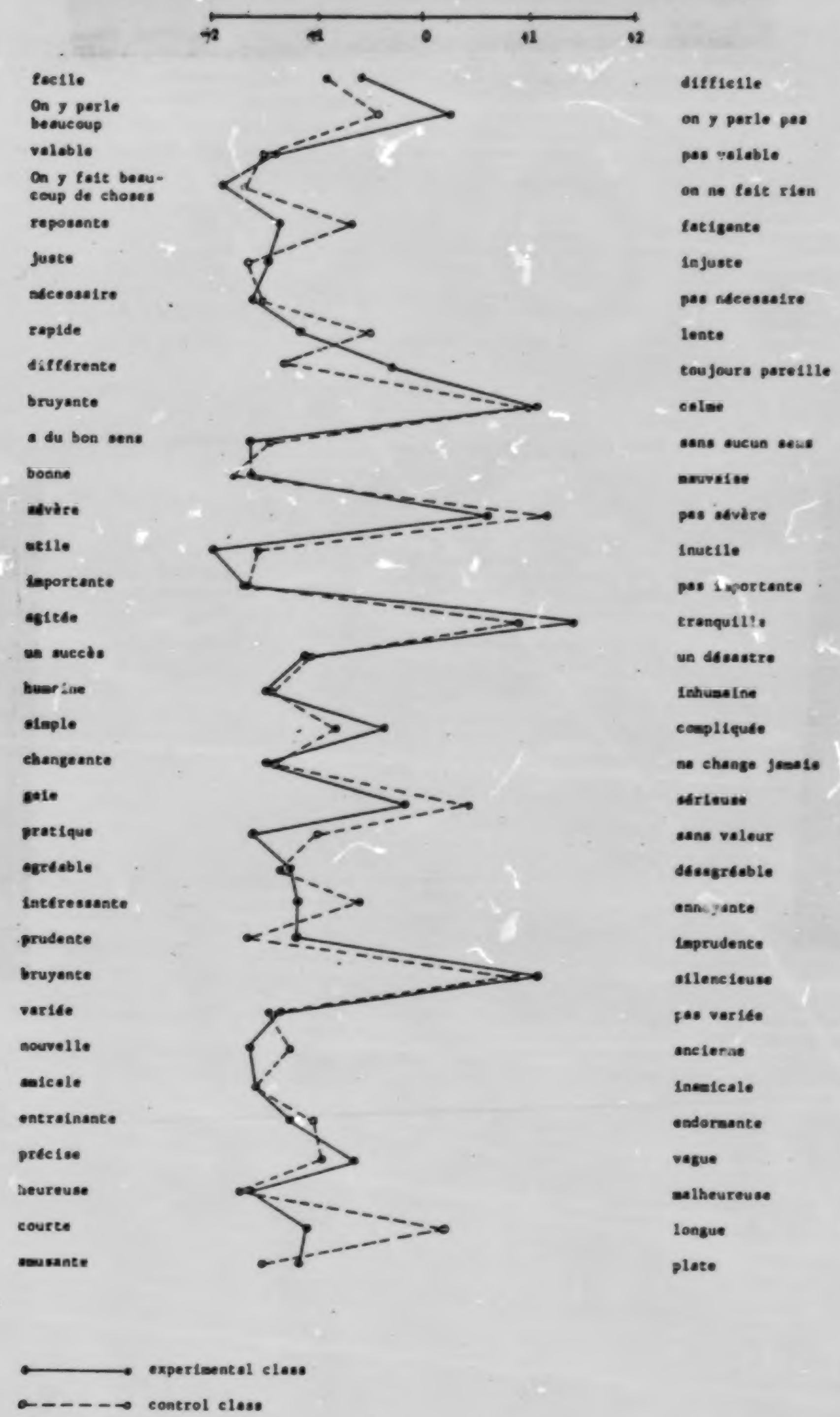


Fig. 21-2

The profiles of the experimental and the control group, rating the concept mathematics lesson against 34 scales in the Sherbrooke Study.

Grade 3

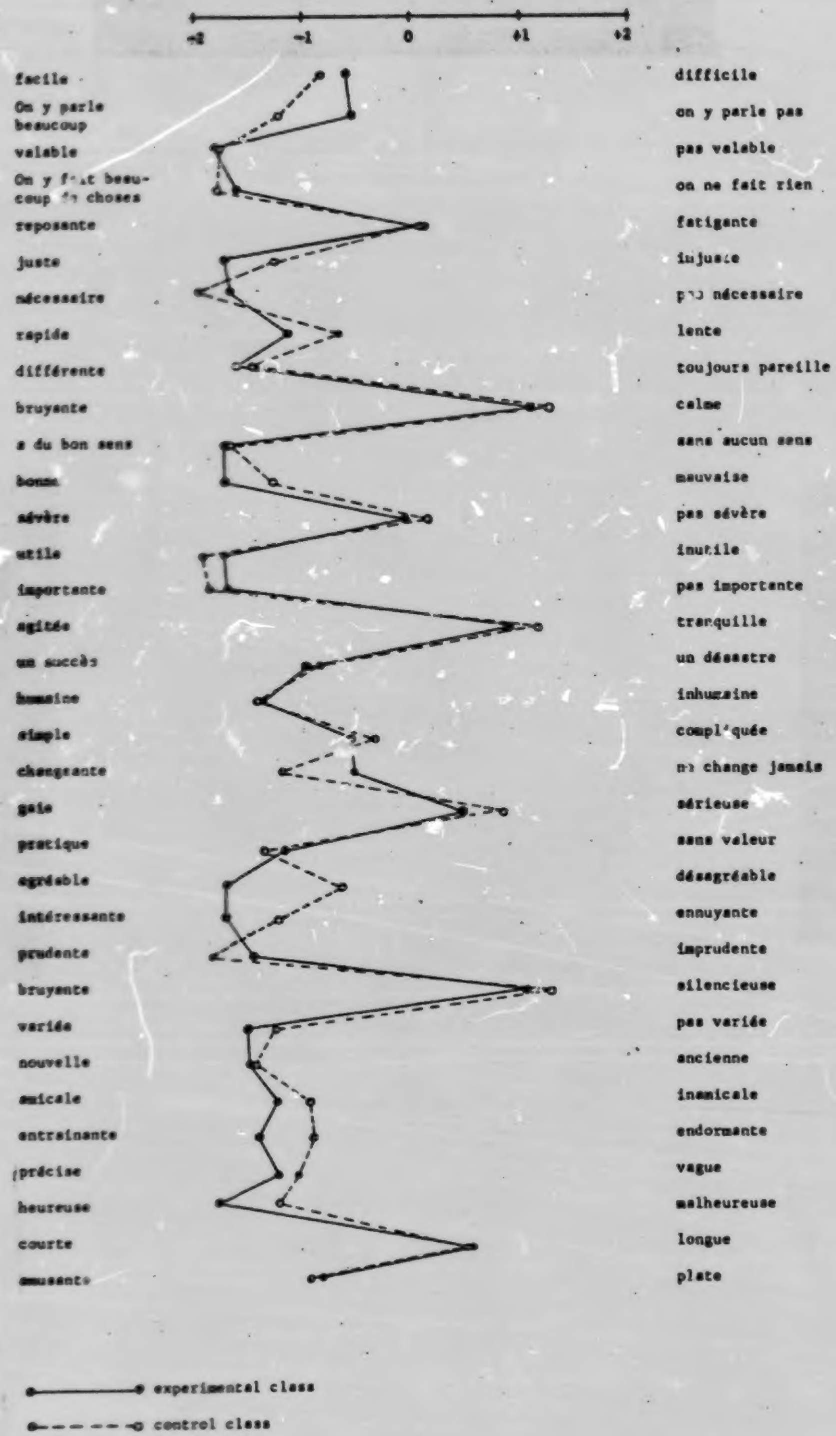


Fig. 21-3

The profiles of the experimental and the control group, rating the concept mathematics lesson against 34 scales in the Sherbrooke Study.

Grade 4

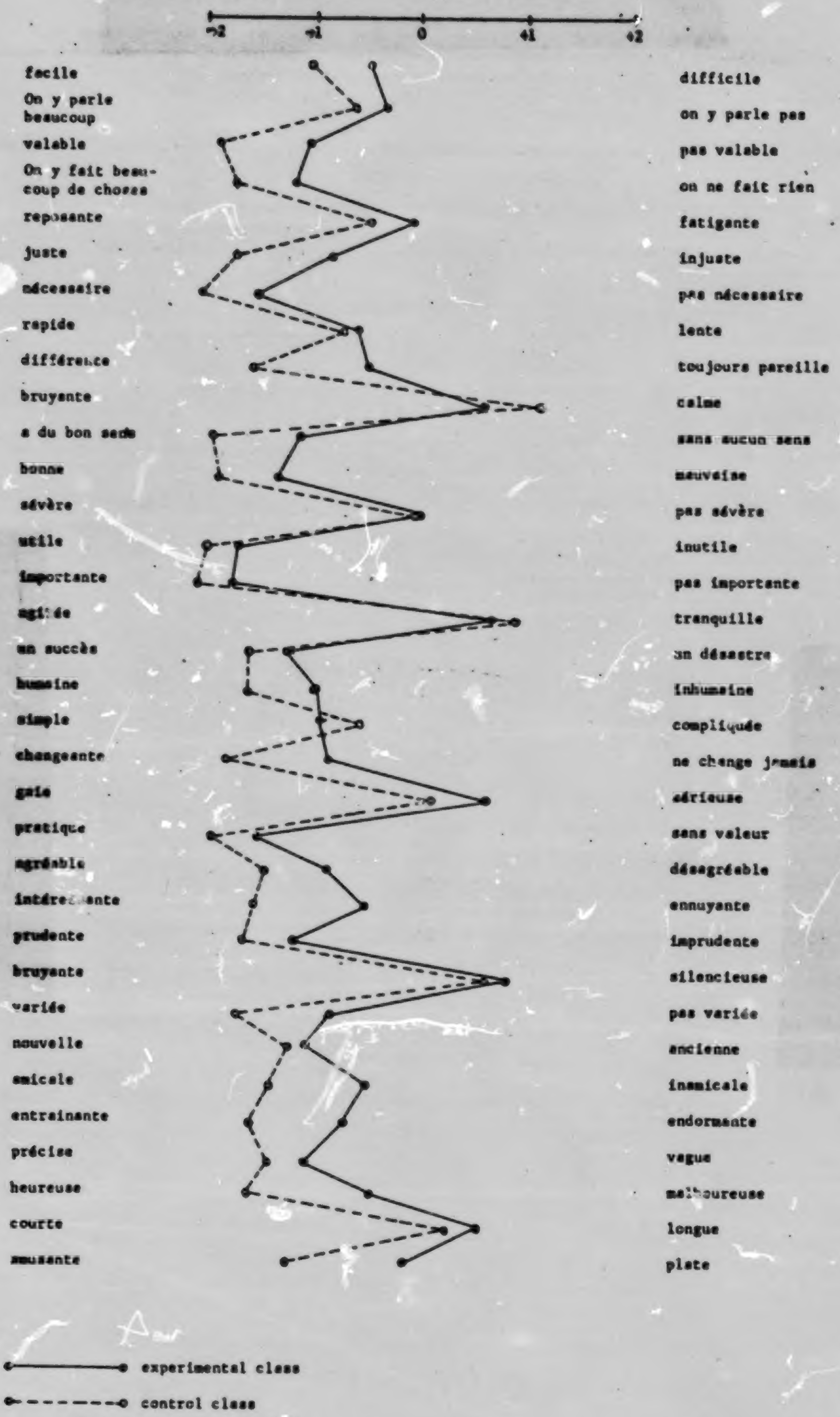


Fig. 21-4

The profiles of the experimental and the control group, rating the concept mathematics lesson against 34 scales in the Sherbrooke Study.

Grade 5

Fig. 21-5

The ratings of the concepts
"Friend" and "Teacher" by class: 4 red and 4 yellow together in the New York Study.

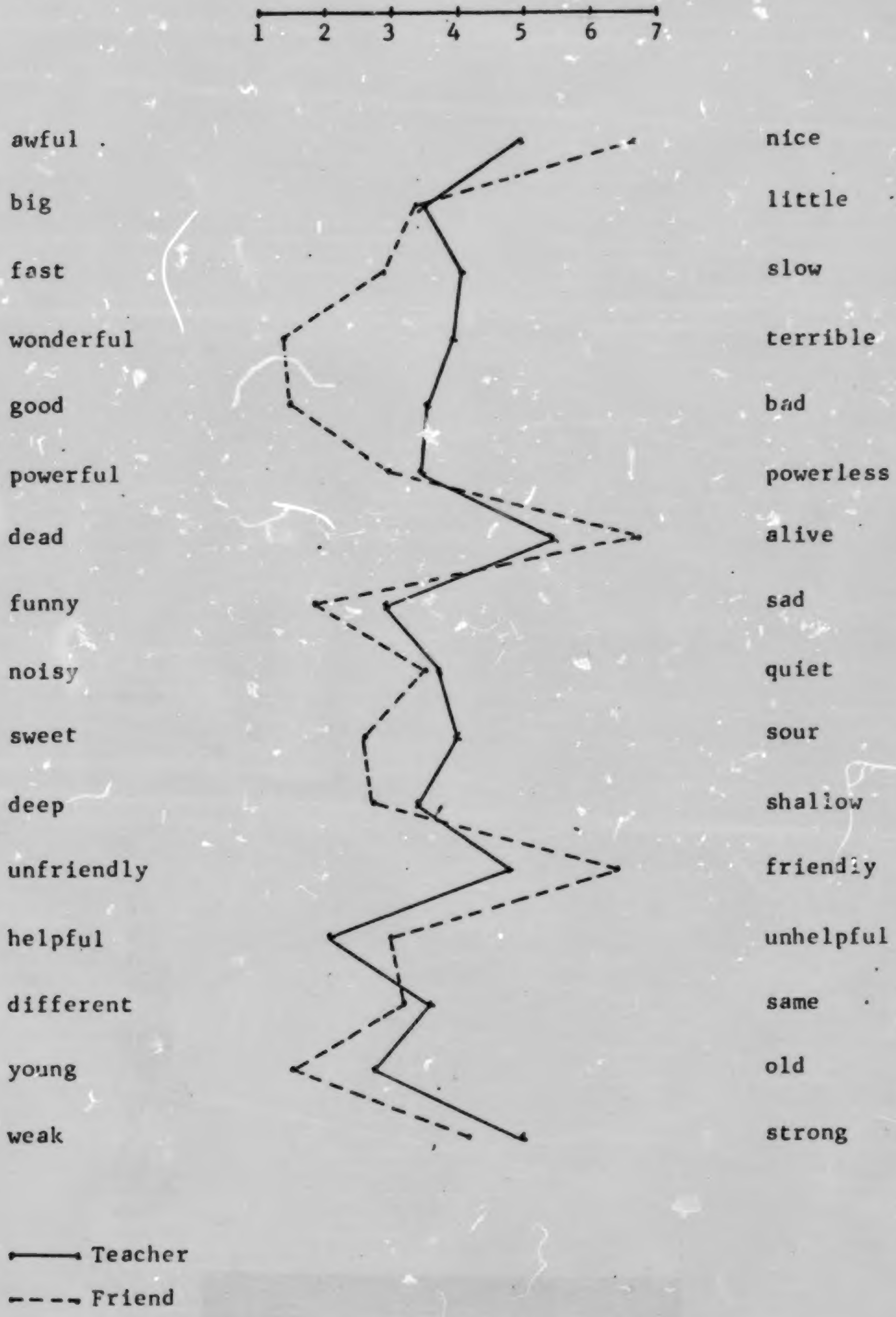


Fig. 21-6

The ratings of the concepts "Friend" and "Teacher" by class: 5 orange and 5 yellow together in the New York Study.

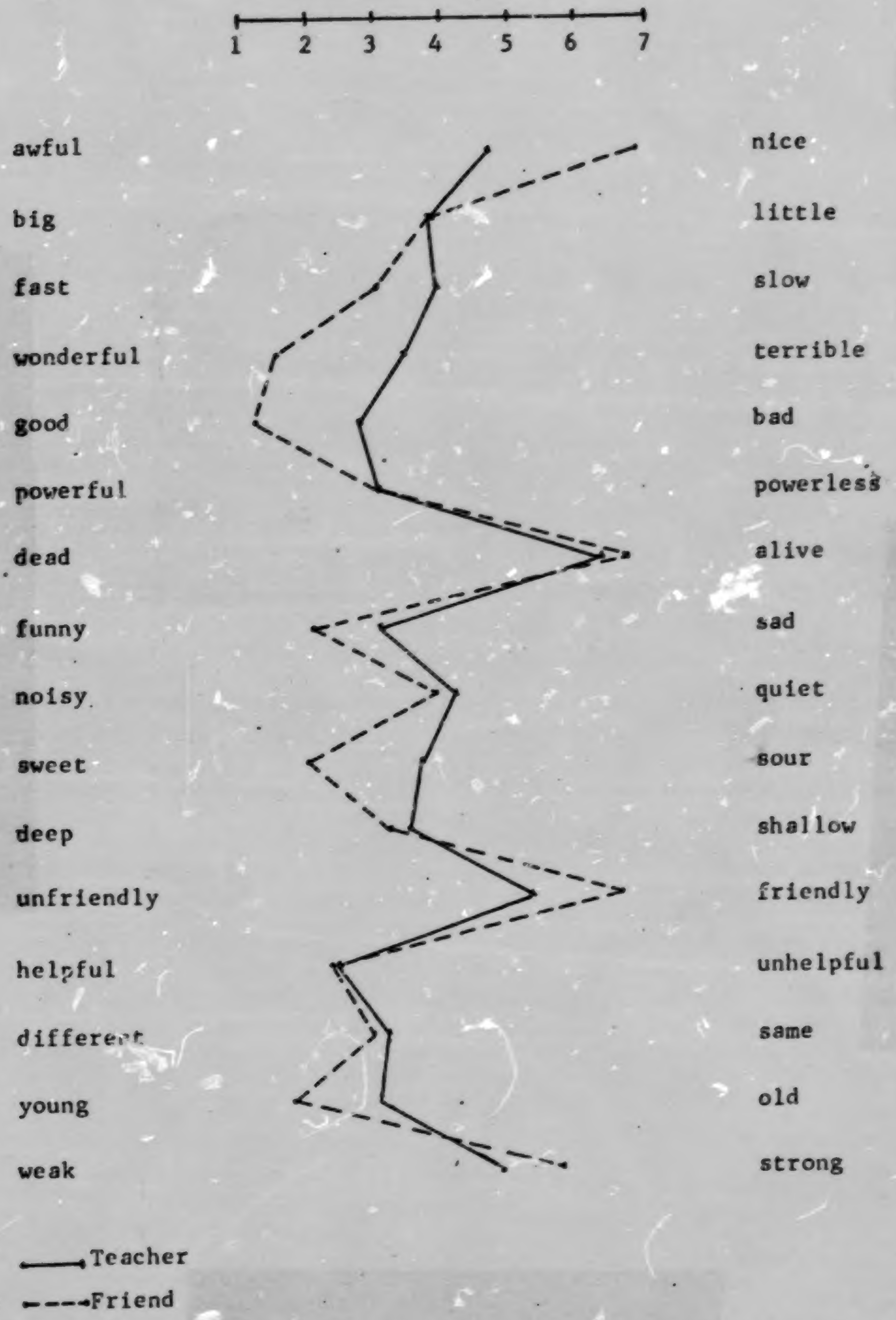


Fig. 21-7

The ratings of the concepts:
"Friend" and "Teacher" by class: 4 red, 4 yellow, 5 orange and 5 yellow together in the New York Study.

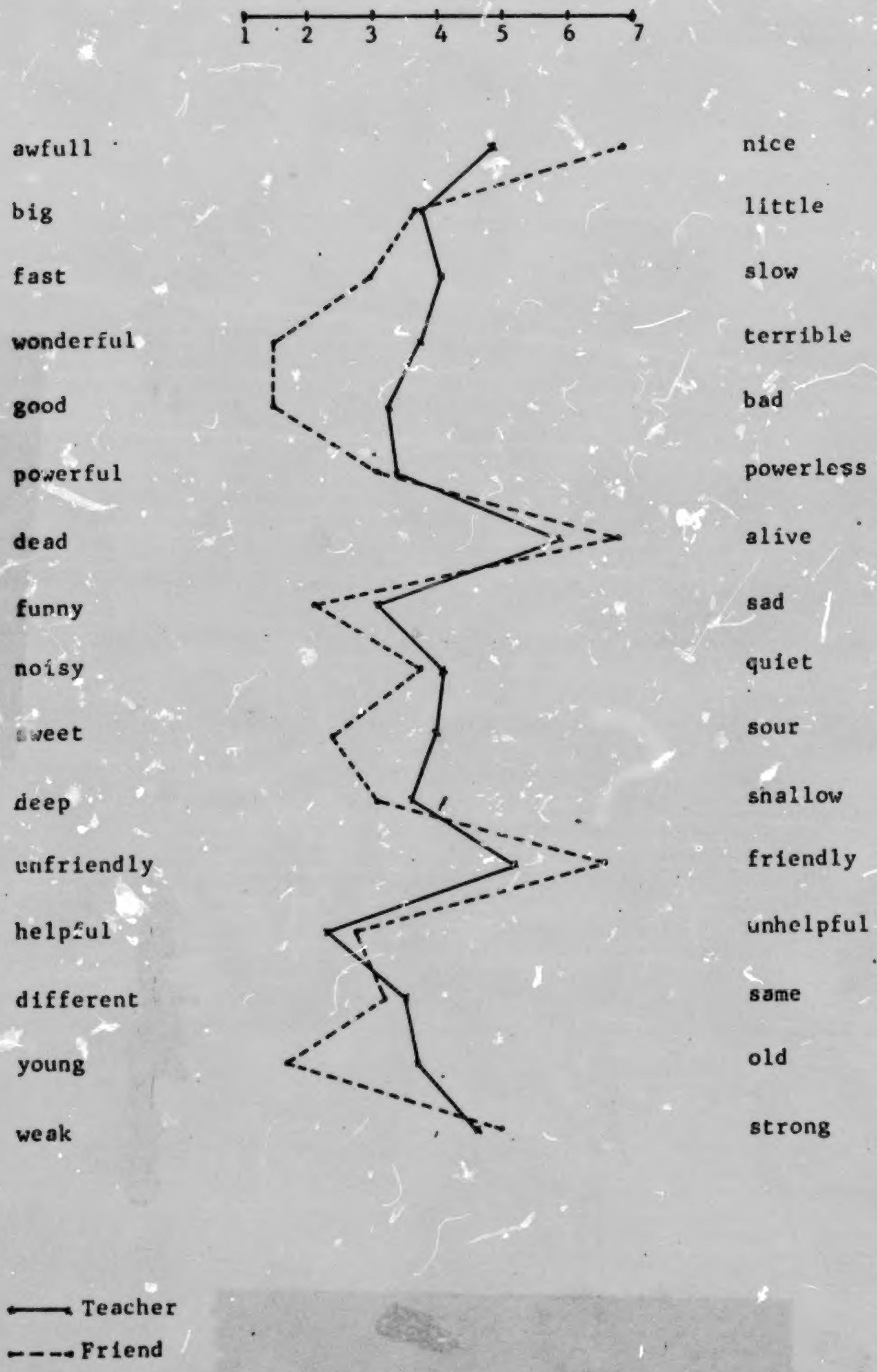


Fig. 21-8

The ratings of the concepts "Math lesson", "School" and "Play" by class: 4 red and 4 yellow together in the New Yord Study.

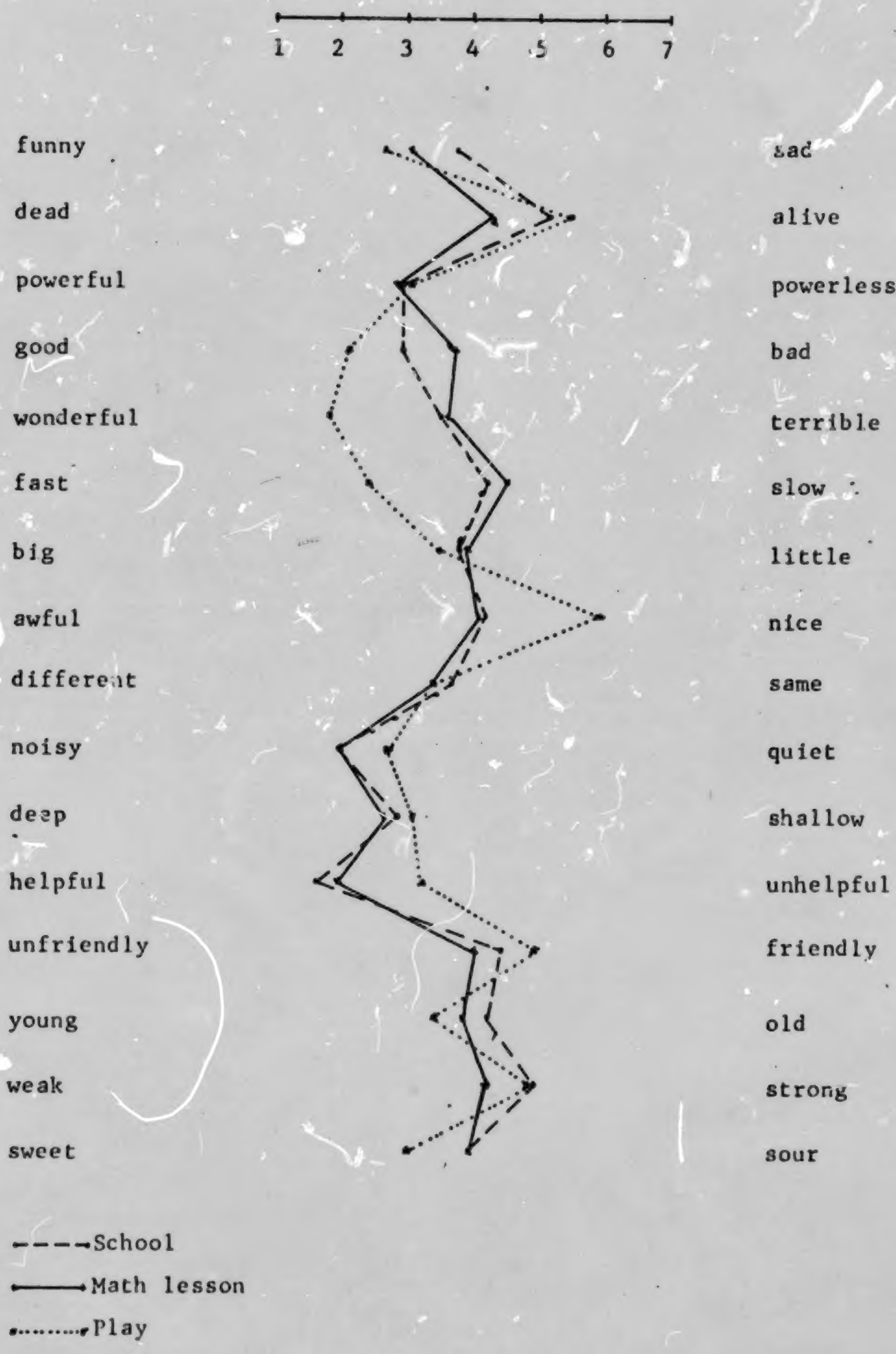


Fig. 21-9

The ratings of the concepts
"Math lesson", "School" and "Play" by class: 5 orange and 5 yellow together in the New York Study.

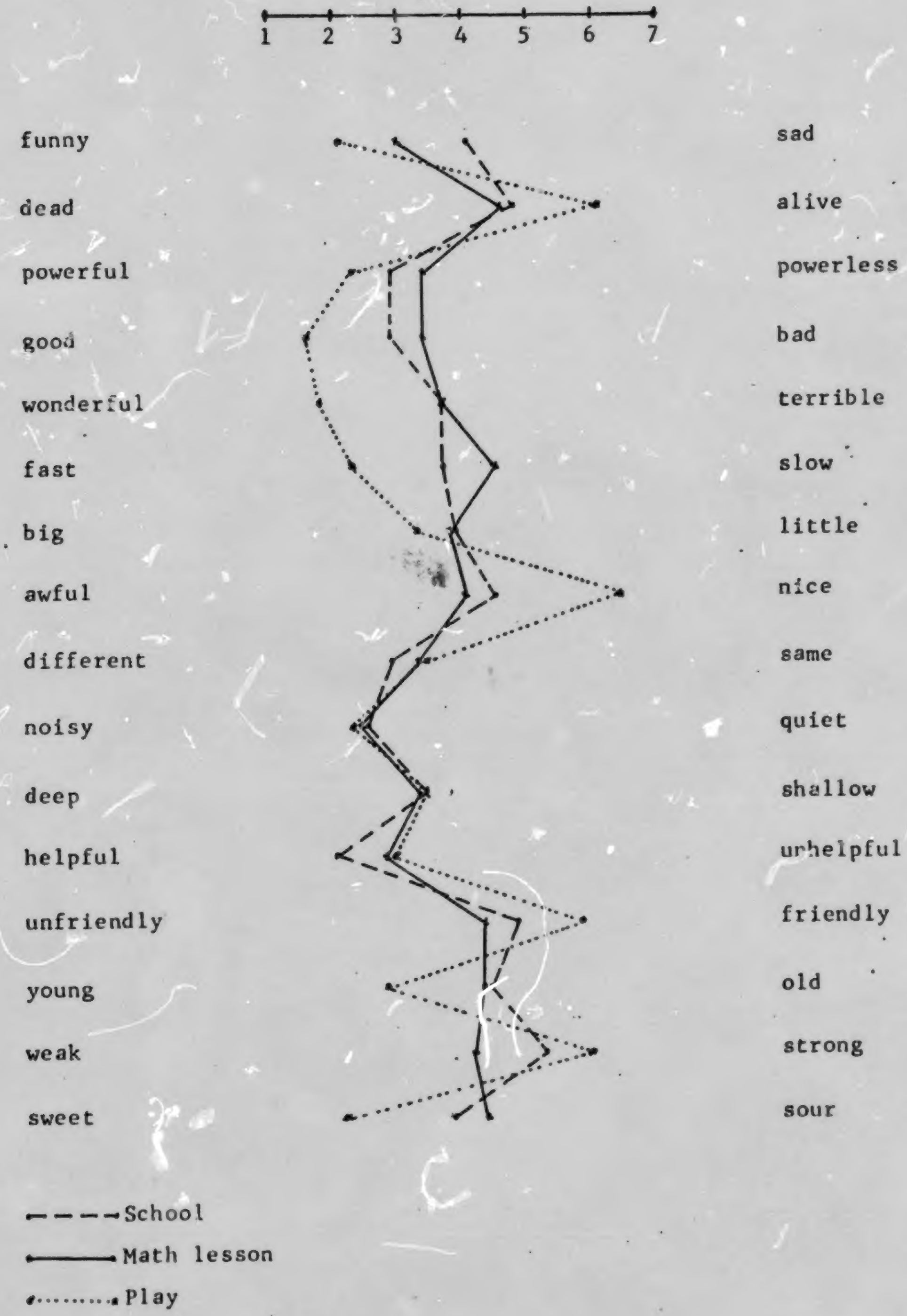


Fig. 21-10
The ratings of the concepts
"Math lesson", "School" and "Play" by class: 4 red, 4 yellow, 5 orange and 5 yellow together in the New York Study.

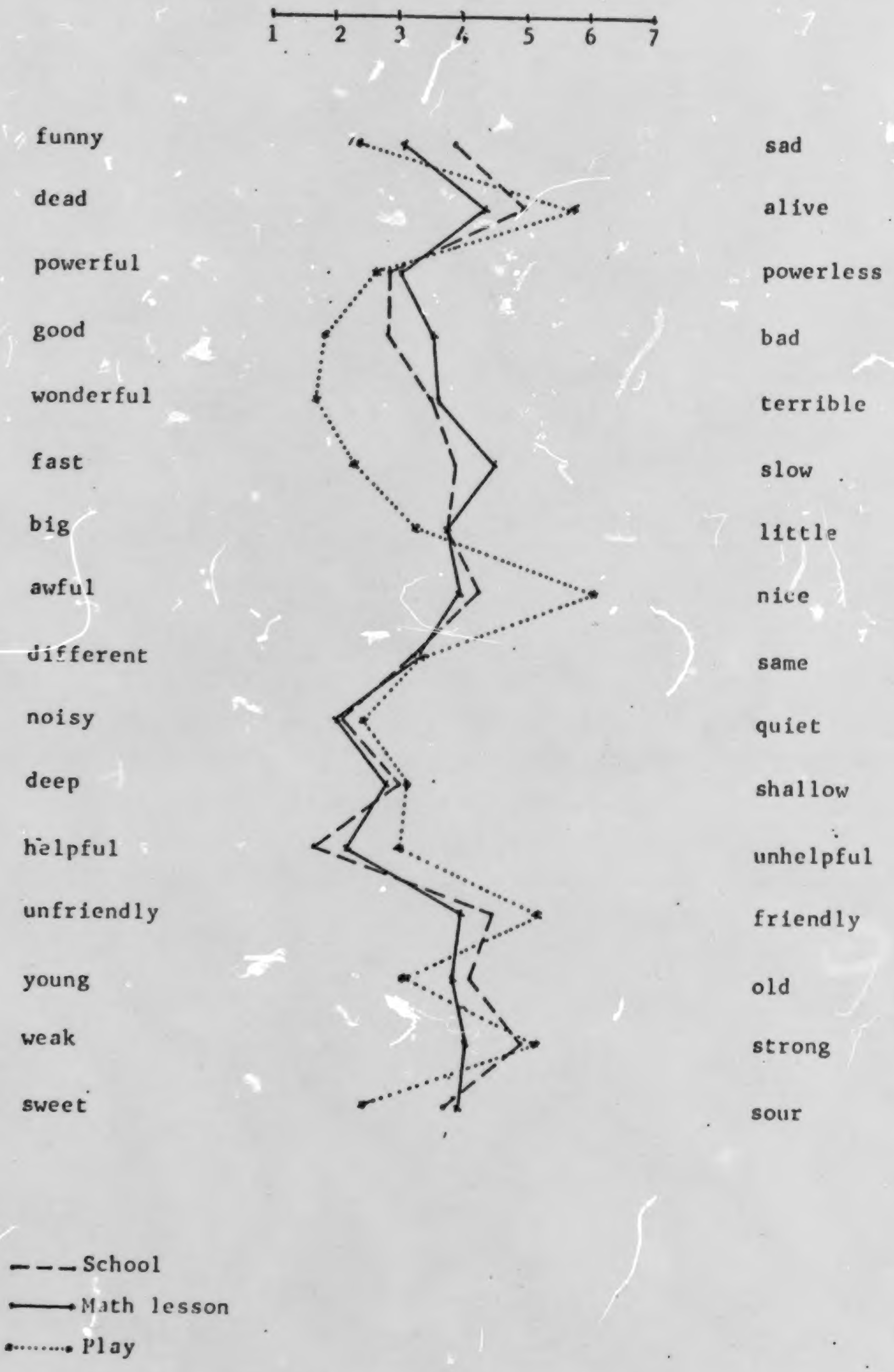


Fig. 21-11

The rating of the concept "School" by the classes: 4 red, 4 yellow, 5 orange and 5 yellow separately in the New York Study.

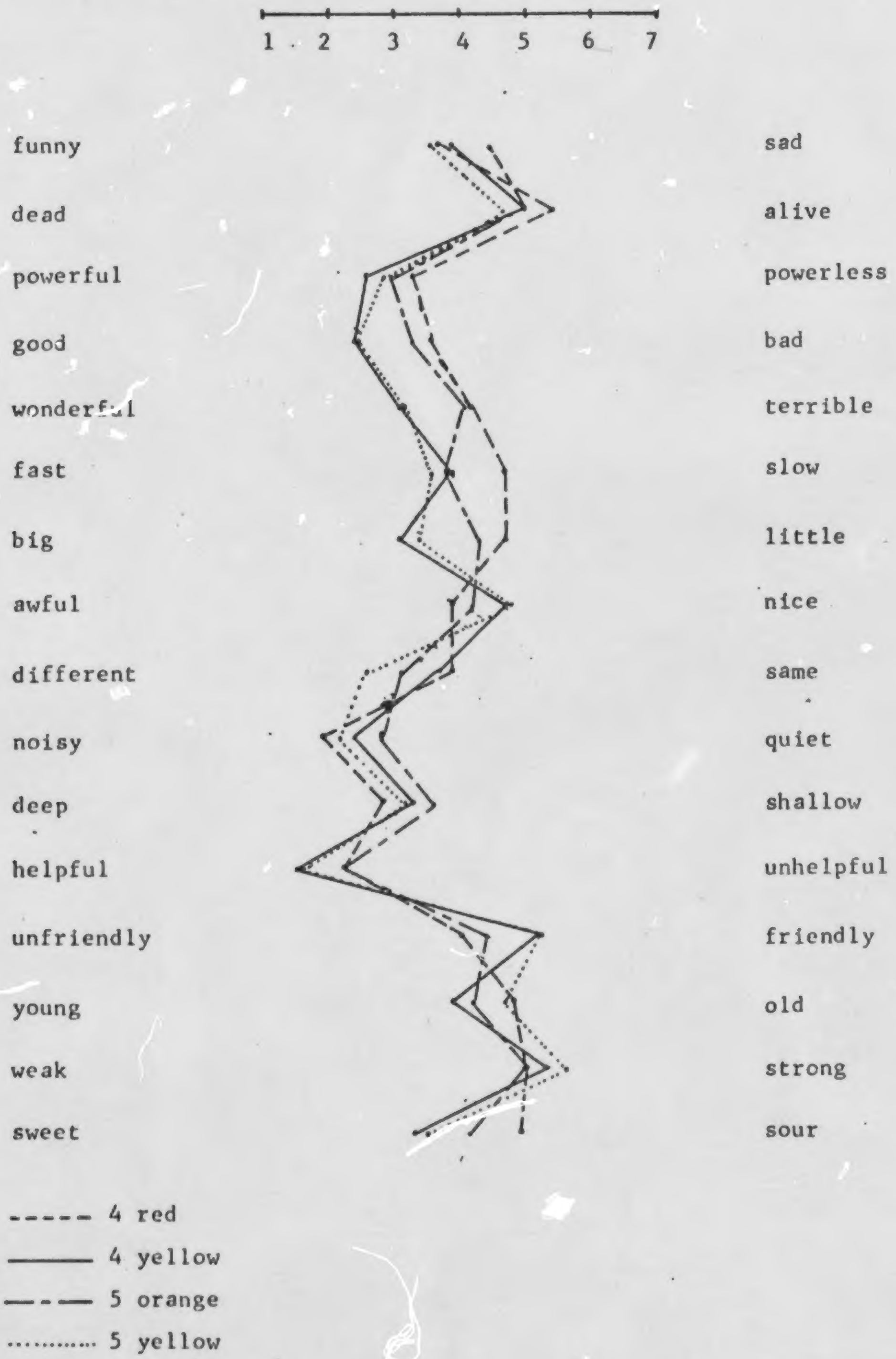


Table 21-4

The seven factors obtained by analysing the ratings of the concept "Math lesson" on 34 bipolar scales, in the Sherbrooke Study.

Grade 3.

Factor I	Factor II	Factor III	Factor IV	Factor V	Factor VI	Factor VII
24 interesting .73	18 kind .78	10 calm .80	25 cautious .75	31 precise .63	19 simple .77	13 lenient .65
30 stimulating .70	7 necessary .69	26 quiet .68	22 useful .58	4 we do a lot .56	17 successful .66	34 amusing .63
29 sociable .65	3 valuable .68	8 fast .62	20 changeable .54	9 different .56	1 easy .45	6 fair .62
23 pleasant .61	14 useful .67	33 short .46	16 calm .52	20 changeable .52	31 precise .36	32 happy .62
5 restful .55	12 good .60	2 quiet .42	15 important .47	33 short .49	30 stimulating .35	1 easy .54
11 meaningful .52	15 important .48	16 calm .42	28 new .46	7 necessary .38	22 useful .32	27 varied .43
27 varied .47	21 serious .48	12 good .39	14 useful .37	15 important .35		12 good .43
34 amusing .42	32 happy .46		17 successful .36	34 amusing .31		8 fast .40
28 new .40	11 meaningful .45		26 quiet .33	10 calm .30		23 pleasant .39
13 lenient .32	17 successful .37		11 meaningful .30			9 different .37

Notes: a) for reasons of space only the left half of the scale is presented here (in cases of negative loadings, the right half).

b) if the same scale adjectives occur under different numbers this is due to translation difficulties.

Table 21-5

The seven factors obtained by analysing the ratings of the concept "Math lesson" on 34 bipolar scales, in the Sherbrooke Study. Grade 4.

Factor I		Factor II		Factor III		Factor IV		Factor V		Factor VI		Factor VII	
32 happy	.85	25 cautious	.81	34 amusing	.70	13 severe	.65	27 varied	.75	9 different	.70	4 we do a lot	.83
23 pleasant	.85	31 precise	.72	33 short	.58	15 unimportant	.63	17 successful	.63	16 calm	.66	8 fast	.43
24 interesting	.76	14 useful	.67	1 easy	.56	18 cruel	.45	10 calm	.63	28 new	.64	28 new	.40
12 good	.69	22 useful	.55	5 restful	.54	8 slow	.41	26 quiet	.47	20 changeable	.56	3 valuable	.34
11 meaningful	.66	7 necessary	.41	19 simple	.54	16 agitated	.36	12 bad	.37	2 quiet	.45	25 cautious	.33
6 fair	.65	19 complex	.39	26 quiet	.47	11 meaningless	.34			22 useful	.44		
30 stimulating	.52	15 important	.34	18 cruel	.43	29 sociable	.33			21 serious	.40		
8 fast	.47			7 necessary	.36	3 valuable	.32			5 restful	.39		
29 sociable	.44					9 always the same	.31			12 good	.35		
34 amusing	.36									26 quiet	.35		

Table 21-6

The seven factors obtained by analysing the ratings of the concept "Math lesson" on 34 bipolar scales, in the Sherbrooke Study.

Grade 5.

Factor I		Factor II		Factor III		Factor IV		Factor V		Factor VI		Factor VII	
15 important	.78	16 agitated	.78	11 meaningful	.73	31 precise	.70	19 simple	.74	2 talkative	.68	29 sociable	.66
9 different	.73	13 severe	.59	12 good	.69	22 useful	.68	8 fast	.73	20 changeable	.63	21 humorous	.63
24 interesting	.69	10 noisy	.59	3 valuable	.65	33 long	.63	6 fair	.53	1 easy	.41	23 pleasant	.61
7 necessary	.66	26 loud	.58	4 we do a lot	.63	6 fair	.35	30 stimulating	.49	26 loud	.39	18 kind	.48
5 restful	.61	7 unnecessary	.52	32 happy	.60	4 we don't do anything	.31	24 interesting	.41	25 cautious	.36	1 easy	.47
34 amusing	.60	19 complex	.34	17 successful	.53	14 useful	.30	12 good	.37			5 restful	.45
27 varied	.60	17 unsuccessful	.31	28 new	.52			33 short	.34			34 amusing	.38
32 happy	.49			18 kind	.50			1 easy	.32			25 cautious	.36
23 pleasant	.48			6 fair	.46							28 old	.34
20 changeable	.47			34 amusing	.35								

Table 21-7

The mean-differences between experimental and control groups on the 7 attitude factors, in the Sherbrooke Study.

Factors	Grades / Statistics	EXP $\bar{X}$	EXP SD	EXP N	CONTROL $\bar{X}$	CONTROL SD	CONTROL N	E X P + CONTL $\bar{X}$	E X P + CONTL SD	E X P + CONTL N	$t^{.05}_{crit}$	t^{1}	$F^{.10}_{crit}$	F
I	5	0.30	1.11	23	-0.43	0.54	21	0.05	0.95	44	2.06	2.67^{*}	2.10	4.21
	4	-0.38	0.46	20	0.23	1.09	18	-0.09	0.88	38	2.07	-2.25^{*}	2.21	5.54
	3	-0.25	0.86	20	0.41	1.15	21	0.09	1.07	41	2.02	-2.08^{*}	2.15	1.82
II	5	0.01	1.22	23	-0.04	0.70	21	-0.01	1.00	44	2.06	0.16	2.10	2.97
	4	0.11	1.33	20	-0.18	0.70	18	-0.03	1.09	[illegible]	2.07	0.80	2.23	3.56
	3	0.10	0.90	20	-0.12	1.07	21	-0.01	1.00	41	2.02	0.71	2.15	1.43
III	5	0.45	1.31	23	-0.33	0.40	21	0.08	1.06	44	2.06	2.57^{*}	2.10	10.67
	4	0.00	0.97	20	-0.07	1.14	18	-0.03	1.05	38	2.02	0.21	2.21	1.39
	3	-0.32	0.79	20	0.11	1.16	21	-0.10	1.02	41	2.02	-1.40	2.15	2.15
IV	5	0.06	1.12	23	-0.13	0.89	21	-0.03	1.02	44	2.02	0.62	2.10	1.59
	4	-0.12	0.92	20	0.17	1.22	18	0.02	1.08	38	2.02	-0.80	2.21	1.76
	3	-0.05	0.72	20	0.08	1.08	21	0.01	0.92	41	2.06	-0.42	2.15	2.26
V	5	0.06	1.18	23	0.07	0.81	21	0.07	1.02	44	2.06	-0.03	2.10	2.13
	4	0.02	0.91	20	-0.03	1.01	18	-0.00	0.96	38	2.02	0.16	2.21	1.23
	3	0.20	0.96	20	-0.11	0.59	21	0.04	0.81	41	2.06	1.23	2.12	2.69
VI	5	0.28	1.09	23	-0.19	0.93	21	0.05	1.04	44	2.02	1.54	2.10	1.36
	4	0.17	1.18	20	0.02	0.87	18	0.10	1.04	38	2.02	0.44	2.23	1.83
	3	0.03	0.90	20	-0.00	1.16	21	0.01	1.04	41	2.02	0.10	2.15	1.68
VII	5	0.17	1.20	23	-0.14	0.86	21	0.02	1.06	44	2.02	1.00	2.10	1.93
	4	-0.04	1.07	20	-0.08	0.76	18	-0.06	0.93	38	2.02	0.15	2.23	1.98
	3	-0.01	0.62	20	-0.39	0.79	21	-0.21	0.74	41	2.02	1.72	2.15	1.60
sum of factors	5	1.33	3.17	23	-1.19	1.63	21	0.13	2.85	44	2.06	3.19^{**}	2.10	3.78
	4	-0.23	2.40	20	0.06	2.76	18	-0.09	2.58	38	2.02	-0.35	2.21	1.32
	3	-0.31	2.32	20	-0.04	2.56	21	-0.17	2.45	41	2.02	-0.35	2.15	1.22

---

* denotes significance at the 5% level, ** at the 1% level.

[1] the formula for t for heterogeneous variances was applied if variances were significantly different at the 10% level F > F crit.[0.1]). In these cases the F's are underlined.

Table 21-8

The mean-differences between girls and boys on the seven attitude factors, in the Sherbrooke Study.

Grades \ SEX		BOYS			GIRLS						
Factors	Statistics	$\bar{X}$	SD	N	$\bar{X}$	SD	N	t^{1}	$t^{.05}_{crit}$	$F^{.10}_{crit}$	F
I	5	0.1	1.12	22	-0.2	0.73	22	2.06	0.79	2.09	2.37
	4	0.2	0.96	24	-0.5	0.42	14	2.10	2.58*	2.42	5.35
	3	0.3	1.00	21	-0.1	1.10	20	2.02	1.24	2.12	1.22
II	5	-0.2	0.84	22	0.1	1.13	22	2.02	-0.94	2.09	1.78
	4	-0.1	1.11	24	0.2	1.02	14	2.02	-0.87	2.42	1.16
	3	0.3	1.26	21	-0.3	0.48	20	2.06	1.82	2.15	6.97
III	5	0.1	1.27	22	0.0	0.80	22	2.06	0.32	2.09	2.51
	4	0.1	1.08	24	-0.2	0.99	14	2.02	0.70	2.42	1.21
	3	-0.0	1.07	21	-0.2	0.96	20	2.02	0.34	2.15	1.24
IV	5	-0.1	0.98	22	0.1	1.04	22	2.02	-0.79	2.09	1.13
	4	-0.0	1.23	24	0.1	0.75	14	2.10	-0.17	2.42	2.68
	3	0.2	1.14	21	-0.2	0.53	20	2.06	1.56	2.15	4.67
V	5	-0.3	0.98	22	0.4	0.93	22	2.02	-2.43*	2.09	1.11
	4	-0.0	1.05	24	0.1	0.77	14	2.02	-0.32	2.42	1.87
	3	0.1	0.67	21	-0.1	0.92	20	2.02	0.70	2.12	1.90
VI	5	0.0	1.13	22	0.1	0.94	22	2.02	-0.18	2.09	1.45
	4	0.2	1.17	24	-0.1	0.72	14	2.02	1.01	2.42	2.65
	3	-0.0	1.05	21	0.1	1.03	20	2.02	-0.30	2.15	1.02
VII	5	-0.4	1.06	22	0.5	0.86	22	2.02	-3.07**	2.09	1.51
	4	-0.0	1.07	24	-0.2	0.63	14	2.10	0.47	2.42	2.89
	3	-0.3	0.80	21	-0.1	0.66	20	2.02	-0.79	2.15	1.47
sum of factors	5	-0.8	1.62	22	1.0	3.45	22	2.06	-2.21*	2.09	4.55
	4	0.3	2.63	24	-0.7	2.38	14	2.02	1.20	2.42	1.22
	3	0.5	2.63	21	-0.9	1.99	20	2.02	1.96	2.15	1.75

The footnotes of the previous Table apply.

Chapter 22

Anxiety - Achievement Motivation - Effects on Personality

22.1 Anxiety

Hypothesis 10: The Dienes method of mathematics learning decreases the students' anxiety.

22.1.1 Theoretical background

Two recent studies in Hungary (Geréb and Szabó, 1965; Geréb, 1970a) show illustratively that anxiety is still a part of school life. Assessed either by recollection of memories about primary and secondary school days or by interviews, observation, and dream analysis, between 23% and 18 to 38% respectively of the individuals studied showed indications that they had to deal with that phenomenon.

Anxiety in general and school related anxiety has gained a large interest in psychological research in the last two decades. Spielberger (1966a) estimates 3,000 publications on this topic between 1950 and 1966, so even for the literature on school-related anxiety it is impossible to give a review here. We may refer the interested reader to a rather exhaustive survey on this by Ruebush (1963) and for anxiety in general to the two volume work, Anxiety: Current Trends in Theory and Research, edited by C.D. Spielberger (1972). Each of the leading contributors to this field seems to approach the problem of anxiety with his own unique theoretical orientation and research methods. It is therefore no wonder that semantic confusion emerges and incompatible and even contradictory results occur. For example, Sarason et al. (1958) found that the less anxious high school and college students show a better performance in learning tasks. Similarly, McCandless and Castaneda (1956), using the children's version of the Taylor Manifest Anxiety Scale (CMAS; Taylor, 1953) on a large school population, found that there is a negative correlation between school achievement and anxiety, especially in those subjects which involve not only mnemonic skills (such as spelling) but also more complicated skills (such as mathematics). There is quite a lot of evidence that a high level of anxiety is detrimental to complex learning or performance (Spence, 1956; Sarason, 1960; Spielberger, 1966; Palermo et al., 1956; Castaneda et al, 1956b). On the other hand, Hughes et al. (1954) did not find any difference between the school achievement of high and low anxiety level students and Pickerel (1958) in some cases found even a better performance in school situations in the high anxiety group. (A study of Alpert and Haber, 1960, shows that this better performance could be the effect of the fear of failure: the achievement of anxious students is higher in highly structured classrooms than in those where organization and structure are more flexible.)

We will refrain from reporting more studies because in order not to be misleading we would have to go into details about underlying theoretical orientations, methods employed, and experimental conditions of each study. Instead, let us gather up our thoughts by considering the points of view on which scientists seem to converge today (after B.N. Philipps et al, 1972).

1. Anxiety is manifested physiologically, phenomenologically, and behaviorally, and discrepancies between these indicators of anxiety may be at least partly attributed to defensiveness.
2. Anxiety has a two-part conceptual status, including what is referred to as "trait," "neurotic," or "chronic" anxiety, and on the other hand what is called "state," "objective," or "situational" anxiety. Trait anxiety is dispositional in nature, is construed to be a proneness to be anxious, is primarily a function of past

experience, and has an internalized (i.e. intrapsychic) locus; while state anxiety is situational in nature, is directly a function of stressful conditions, and has a contemporary locus. Moreover, these two variables interact in a manner such that anxiety proneness influences the extent of the anxiety reaction.

3. Anxiety is elicited by psychological stress, and stress is reflected in threatened deprivation of an anticipated satisfaction. In addition, uncertainty associated with external (i.e., environmental) and internal (i.e., cognitive) factors is a key ingredient in whether an anxiety reaction to stressful conditions occurs.

4. Anxiety usually occurs as a response to stress in conjunction with other affects, defensiveness, other coping reactions, etc.

5. The consequences of anxiety are usually negative, interfering and debilitating in nature, although all consequences depend on the demands and requirements of situations. Therefore, the consequences are complex and interactional in nature - and, as a result, are sometimes facilitating, helpful, and enhancing to adaptation, performance, and adjustment.

Furthermore, anxiety may be experienced on different levels, which are developmentally determined: on the senso-motoric level in a motoric, sensory, and physiological way; on the perceptual level in terms of imagery, fantasies and the like; and on a conceptual level in terms of negative valuations of self, regressive behaviour, and other defensive tendencies.

The points stated above show that the concept of anxiety as a topic of research is very complex. Before we can fix our standpoint in this net of relationships, let us take a closer look at possible factors which seem to be responsible for anxiety (cf. Phillips, 1972). We will neglect here the variables which relate to or cause situational anxiety, since our concern is to trait anxiety, which may be seen as the outcome of the individual's integration or ability to cope with situational factors, most generally stress. Some of the important factors in the origin of anxiety are:

1. Environmental factors:
    a. Specific past experiences, especially preschool and early school experiences, in which important motives have been thwarted.
    b. Racial-ethnic minority status and/or lower social class status.
    c. Specific aspects of the parent-child relationship.
    d. Patterns of socialization of threat reactions.
2. Organismic factors:
    a. Sex and sex role identification.
    b. Birth order.
    c. General beliefs about the environment, and one's relationship to it.
    d. Counterharm resources of the individual (e.g. his intellectual abilities).

Measurement of anxiety in our study is concerned with anxiety caused by school experiences only, the other factors are neglected, excluding sex and social class status, which were partly controlled and supposed not to interfere as co-variables.

If anxiety is such a complex and rather ambiguous concept in research, what can we assume to measure? In fact, in general, different measures of anxiety seem to measure different components of it: the relationship between pencil and paper measures of anxiety and physiological measures is usually very low or even controversial; there is also, generally, a low correlation between clinically rated anxiety (by a therapist), self-rated anxiety and psychometric anxiety (Spielberger, 1966a; p. 233; Cattell and Scheier, 1961). For our study, we are measuring that aspect of anxiety which can be revealed by the kind of questionnaires employed (see description below). It is an aspect of trait anxiety, i.e. proneness to be anxious in a variety of different situations with emphasis on school situations. We have to keep in mind that although the items of the questionnaires may refer to physiological as well as to imaginative behaviour and defense mechanisms, only that behaviour which passes through the sieve of the child's introspection can be covered by the measurement and enter into the results. We have included anxiety as a variable in our study despite its complexity and its difficulty to measure because school experiences are a strong but unnecessary source of children's anxiety, with far-reaching consequences. The following list shows some of these consequences which have been researched:

1. Some proximal consequences:
    a. Cautiousness, perseveration, rigidity, stereotyped thinking.
    b. Dependency, direction-seeking, conforming tendencies.
    c. Reduced responsiveness to the environment.
    d. Interference with a variety of cognitive and mediational processes.
    e. Increased drive or motivational level.
2. Some distal consequences (= consequences of the above):
    a. Deterioration in complex intellectual, problem solving, achievement, and learning activities.
    b. Increased responsiveness to reinforcement.
    c. Increased susceptibility to persuasion, and the behaviour of models.
    d. Preoccupation with demands of self rather than the demands of learning situations, leading to reduced incidental learning, etc.
    e. Increased isolation from others, with heightened affiliation needs.

f. Enhanced learning of certain types of tasks.

With respect to scholastic behaviour on which research has concentrated (see Sarason, 1960; Alpert and Haber, 1960; Spielberger, 1966a), there is general agreement that the effects of anxiety are largely maladaptive and debilitating. It is not worthwhile to report single studies here since relationships between anxiety, intelligence, and achievement are intricate, and only a few studies employing multivariate procedures have ever been performed. Even fewer research efforts have been conducted on the problem we are concerned with, namely how different school-learning environments could minimize anxiety or its undesirable effects. Important results to consider in this framework have been presented by Biggs (1962; Anxiety, Motivation, and Primary School Mathematics). In measurements of general anxiety and in ten anxiety questions similar to Sarason's Anxiety Scales, he could find no evidence of anxiety as a consequence of the arithmetic teaching method, but could demonstrate a relationship between number anxiety (questions relevant to arithmetic) and traditional teaching methods.

Our own work in this area uncovers the idea that independently of what research studies may reveal, any new teaching method has an important humanizing task, to decrease the students' anxiety level. To reach this objective, the Dienes mathematics programme has tried to reduce stress in the classroom by creating and encouraging an environment in which there is no emphasis on grades, no punishment, no pushing from behind, but instead interesting problems, the right to make mistakes, and the right to be different from others (see Table 13-1).

22.1.2 Budapest study

METHOD

To examine whether the OPI project did not fail in this direction we had two main measuring instruments:

(i) the Children's Manifest Anxiety Scale (CMAS; Castaneda, 1956, 1961; Castaneda et al., 1956a)

(ii) the Test Anxiety Questionnaire (TAQ; Mandler and Sarason, 1952).

We had two reasons, one theoretical and one practical, for choosing the first one as a measurement of the students' anxiety. The theoretical reason was that while the CMAS, being the children's version of Taylor's Manifest Anxiety Scale (MAS; 1951; 1953), which uses a range of questions chosen from the Minnesota Multiphasic Inventory (MMPI), is a general measurement of anxiety, the TAQ was specially designed to indicate the extent to which a subject experiences anxiety before, during, and after three "typical" kinds of testing situations. However, testing is not a typical situation in Hungarian schools.

The practical reason was that the CMAS had a well-developed Hungarian version (Forrai, 1968), and that it was better suited to the subjects' age group - in our case, grades 3, 4, and 5.

Unfortunately, because of this decision we had to face some other difficulties in explaining the interaction between anxiety and achievement motivation (see hypothesis 11).

TREATMENT OF DATA

The Mann-Whitney test was used to compare the CMAS scores of experimental and control classes. In addition, scores were grouped into low, middle, and high anxiety levels according to the Hungarian CMAS standards, and the percentages of students in the different levels were compared in the experimental and control classes.

EXPECTATION

Lower CMAS scores in the experimental classes than in the control classes would show decreased anxiety, thus supporting hypothesis 10.

RESULTS

As expected, a significantly lower average CMAS score was found in the grade 4 experimental class (12.7 vs. 17.2), though the differences in the other grades were not significant (18.6 vs. 20.8 in grade 3, and 21.9 vs. 20.5 in grade 5). (These averages are quite like those of Castaneda et al., 1956, and Forrai, 1968). These results thus partially support hypothesis 10.

In Table 22-1 we grouped the students into high, middle and low anxiety level groups. It is interesting to see that in grade 4 and especially in grade 5 the anxiety level of the students is more polarized in the experimental classes than in the control classes. If this measurement reflects a real difference, it could be an important warning to those who are responsible for continuing the experimental teaching method to take more care of those students with less ability and greater anxiety.

Table 22-1. The percentages of the students in the low, middle and high anxiety group in the Budapest Study.

Grades	Anxiety Level	Low (0-15)	Middle (16-26)	High (27-42)
3	Exptl	11	58	31
	Control	28	48	24
4	Exptl	63	28	9
	Control	45	48	7
5	Exptl	31	44	25
	Control	18	73	9

22.1.3 Sherbrooke study

METHOD

Instrument: In the Sherbrooke study, anxiety was measured with Demangeon's *Epreuves d'anxiété* (1966a). in which some items were slightly changed in order to adapt it for a Quebec population. Demangeon's test was used mainly because it was in French and it contained a rather large amount of school-related items. Different forms are used for girls (EA F6/7) and boys (EA G 6/7). Both are designed for age groups 10 to 12 years corresponding to grades 6 and 7 in the French school system. The test consists of two parts:*

The first part "Phrases à compléter" is made up of 40 sentences to be completed, where about 10 of those are related to school situations; the others refer to anxiety which could be aroused in special situations (confrontations with explosions, thunder, animals, sleeping alone in a dark room, etc.). For each sentence, the subject has to choose one of four possibilities to complete it. For example:

Many times I imagined...

that I was a champion	O
that I would visit Europe	P
that there is a burglar in my room	Q
that I had a car	R

A score of one point is given for every anxiety-related complement chosen; the range is therefore 0 - 40 points on this part. The second part "Questionnaire" is made up of 45 items out of which approximately 30 are related to school anxiety.

---

* According to Demangeon the two parts of the test are uncorrelated.

They are given in the form of statements to which the subject has to add a word from one out of four possible choices: "Does it happen that you have bad dreams concerning school? Often; Sometimes; Rarely; Never." For scoring, weights from 3 (Often) to 0 (Never) were accorded so that an individual may get a score between 0 and 135 in this part of the test. Items not related to school refer generally to "worrying, sadness, bad dreams, etc." The two forms for girls and boys are only different in 5 items, the questions about weeping and blushing for example, being replaced by others. Measurements about validity and reliability reported by the author (1966b,1971) seem unsatisfactory when evaluated by the American standards for test development. Demangeon found for the questionnaire reliability coefficients of .78 for stability based on two parallel forms, and of .74 and .81 respectively for homogeneity based on the split half method for the two forms. For the first part (sentences), he reports a stability coefficient of .45. Validity was measured mainly by comparing low anxious and high anxious groups, as determined by their teachers primarily. Consequently, since the anxiety observed and the anxiety experienced (if we assume that the subjects' responses were honest) may be very different things, discriminating power was found in school-related items of the questionnaire only.*

However, the items seemed to us to form a fairly representative pool of how anxiety could occur, because physiological, behavioural, imaginary and neurotic (obsession) indices were equally represented in the questionnaire.

Subjects: The test was given to all subjects in grades 3, 4, and 5 and data from all 123 permanent students were obtained.

---

* An item has discriminating power if it discriminates between 2 groups of subjects. In the case of an anxiety test an item has discriminating power concerning school achievement if, for example, students with a high achievement in school have a low score on this item but students with a low achievement a high score.

Raw scores were counted for each individual for both parts of the test. Mean differences between girls and boys were established for the three grades separately, which revealed significant differences only in the first part of the test (sentences) (see Tables 22-2 and 22-3).

Because of this, experimental-control differences were counted separately for boys and girls only for the first part of the test. The results (see Tables 22-4 and 22-5) did not support our hypothesis because they did not reveal any significant differences in the anxiety of the children in experimental and control classes.

22.2 Achievement Motivation

Hypothesis 11: The Dienes method of mathematics learning increases the students' achievement motivation.

22.2.1 Theoretical background

Motivation[*] is one of the most important factors in learning (Melton, 1952). It is a well-proven fact that otherwise equal but more highly motivated pupils do better in every school level and later in their occupations (Minor and Neel, 1958; Weiss et al., 1959; Shaw, 1961; Uhlinger and Stephens, 1960; Meyer et al., 1965). There is also some evidence that motivation, teaching method and academic

---

[*] The variable "motivation" was examined in the Budapest Study only.

Table 22-2. Differences between boys and girls in Demangeon's Anxiety Test. From "Epreuves d'Anxiété", Part II of the Questionnaire in the Sherbrooke Study.

SEX / Statistics / GRADES	MALE			FEMALE						
	$\bar{X}$	SD	N	$\bar{X}$	SD	N	t	$t_{crit}^{.10}$	F	$F_{crit}^{.10}$
5	49.55	13.60	22	52.23	15.25	22	-0.62	1.67	1.26	2.08
4	43.00	17.99	24	41.93	13.84	14	0.21	1.68	1.69	2.42
3	49.33	19.64	21	39.80	18.32	20	1.61	1.73	1.15	2.15
$\sum^1$	1.16	17.43	67	-1.39	16.32	56	0.83	1.64	1.14	1.55

The method of combining scores is described in the footnote of Table 15-3.

Table 22-3

Differences between boys and girls in Demangeon's Anxiety test: "Epreuves d'anxiété", Part I: "Phrases à compléter" . Sherbrooke Study.

SEX	MALE			FEMALE			
Statis-tics GRADES	$\bar{X}$	SD	N	$\bar{X}$	SD	N	t
5	5.68	2.95	25	8.73	4.56	30	2.99**
4	5.46	3.70	28	7.80	3.60	25	2.33*
3	5.80	3.46	30	7.71	4.45	28	1.82

* significant at .05 percent level

** significant at .01 percent level

Table 22-4

The mean-differences between experimental and control classes in Demangeon's Anxiety test, part II: "Questionnaire". Sherbrooke Study.

GROUPS	EXP			CONTROL			E X P + CONTL				
Statistics / GRADES	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	t*	F*
5	49.83	15.72	23	52.05	12.96	21	50.89	14.51	44	-0.51	1.47
4	44.90	15.95	20	40.06	16.92	18	42.61	16.59	38	0.91	1.13
3	46.95	19.02	20	42.52	19.89	21	44.68	19.59	41	0.73	1.09
Σ^1	1.06	16.98	63	-1.11	16.91	60	0.00	16.98	123	0.71	1.01

* None of them is significant

1 The method of combining scores is described in the footnote of Table 15-3.

Table 22-5. The mean differences between experimental and control classes for boys and girls separately in Demangeon's Anxiety Test.
From Part I, Phrases a completer, in the Sherbrooke Study.

Groups / Grades	Statistics	EXP			CONTROL			EXP + CONTL				
		$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	t[1] *	F
Boys	5	4.55	2.68	11	6.55	3.03	11	5.55	3.03	22	-1.64	1.28
	4	6.55	2.31	11	5.69	4.39	13	6.08	3.62	24	0.56	3.61
	3	7.10	3.96	10	4.82	2.66	11	5.90	3.53	21	1.53	2.22
	Σ[2]	0.19	3.17	32	-0.17	3.60	35	0.00	3.41	67	0.44	1.29
Girls	5	9.25	4.87	12	8.60	2.73	10	8.95	4.05	22	0.36	3.18
	4	8.67	3.94	9	6.60	3.72	5	7.93	3.99	14	0.97	1.12
	3	8.10	3.48	10	7.00	4.27	10	7.55	3.93	20	0.63	1.51
	Σ[2]	0.51	4.20	31	-0.63	3.63	25	0.00	3.99	56	1.08	1.34

* None of them is significant

1 the formula for t for heterogeneous variances was applied if variances were significantly different at the 10 percent level ($F > F^{.10}_{crit}$; underlined).

2 the method of combining scores is described in the footnote of Table 15-3.

achievement have some inter-relation (McKeachie, 1961).

There exist some studies showing that motivation is especially important in a less structured way of learning mathematics, being more related to performance in "unscheduled" than in "scheduled" periods (Fig. 22-1). Motivational factors have an important role in insightful problem-solving (See Chapter 23), so essential in mathematics (Birch, 1945). Achievement motivation was included in the present study, partly because it was supposed that the experimental teaching method has an effect on it and partly because it has a standard assessment device based on a well-developed theory.

MOTIVATION TO ACHIEVE SUCCESS (M_s) AND MOTIVATION TO AVOID FAILURE (M_{af}): THE TWO COMPONENTS OF ACHIEVEMENT-RELATED MOTIVATION.

According to the achievement motivation theory developed by McClelland, Atkinson and their co-workers (McClelland et al., 1953; Atkinson, 1958, 1964; Atkinson and Feather, 1966), the achievement - related motivation (or, as Atkinson and O'Connor (1966) expressed it, "resultant achievement motivation") refers to the resultant of two opposite vectors: the motivation to achieve success (M_s) or need to achieve (n Ach) and the motivation to avoid failure (M_{af}). (In our case we supposed that $M_s > M_{af}$, such that the students' achievement-related motivation had a positive value in both experimental and control classes.)

We expected that the composite method of mathematics teaching would lower the students' motivation to avoid failure and increase their motivation to achieve success, which together could result in quite a large change in the resultant achievement motivation. Let us look at each part of the hypothesis separately.

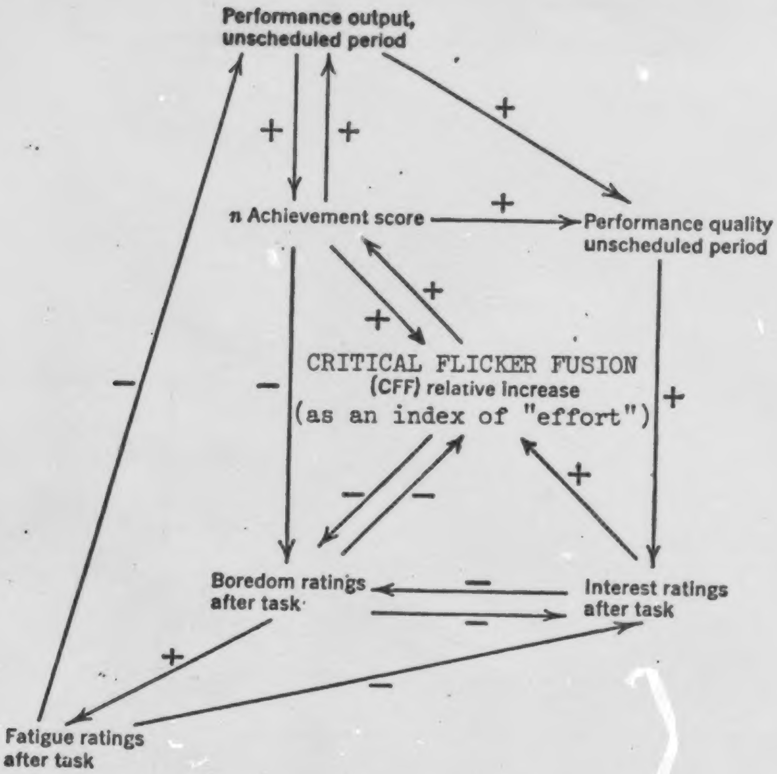

Performance output, unscheduled period
n Achievement score
Performance quality unscheduled period
CRITICAL FLICKER FUSION (CFF) relative increase (as an index of "effort")
Boredom ratings after task
Interest ratings after task
Fatigue ratings after task
+
+
+
+
+
+
+
+
+
−
−
−
−
−
−
+
−

Fig.22-1 Interrelations of variables which are significant at approximately, or better than, the .05 level (high school sample, N = 38). After Wendt, 1955.

628

Motivation to avoid failure (M_{af}):

Heckhausen (1963) and Meyer et al. (1965) pointed out that a predominantly success-oriented achievement motivation alone has a favourable effect on different kinds of school performance. The OPI Mathematics Project has tried to organize a success-oriented educational environment giving the pupils liberty to make mistakes. Under this condition errors often throw more light on the essence of the problem than an elaborate explanation. The work is matched to the pupils ability, so even the "slowest" students can have success experiences.

The traditional way of mathematics teaching could be labelled as a success-failure oriental situation where the failure of a fairly large fraction of students is quite a dominant factor. Because of this we would expect a stronger motivation to avoid failure (M_{af}) in the control classes than in the experimental classes.

It is usually supposed (e.g. Atkinson and Litwin, 1960; Atkinson, 1964) that the Test Anxiety Questionnaire (TAQ) provides a measure of the M_{af}. However, to examine the 10th hypothesis we used the Children's Manifest Anxiety Scale (CMAS; Castaneda et al., 1956a) which characterizes children's general anxiety. As there is a connection between the two types of measurements of anxiety (e.g., Raphelson (1957) got a 0.53 correlation, significant at the 0.01 level between the two tests) we can suppose that the results of the CMAS are indicators of M_{af}. In this case a result supporting the 10th hypothesis (less anxiety in the experimental classes) would support our expectation of a weaker motivation to avoid failure (M_{af}), which could be one factor in the supposed bigger achievement motivation in experimental classes. (For some other aspects of the inter-relation between achievement, anxiety, and learning conditions, see e.g. Sarason et al., 1952.)

Motivation to achieve success (M_s):

As outlined above, the composite method may help to develop in the student a bigger need to achieve. It gives a variety of fields in which the student can create (and so achieve) something and feel the success that accompanies achievement.

As independence and achievement are closely related (Vernon, 1969), the Project emphasises letting the student achieve his own through his own independent actions.

According to our observations the experimental condition encouraged what Atkinson and Feather (1966) named an "achievement-oriented personality"; one that likes challenge and the sense of having done something well, that does not rest content with continual mastery of old familiar tasks when new possibilities open, that approaches quite ambiguous tasks with greater confidence, that can construct a world of interesting challenges, that prefers tasks which require personal initiative and inventiveness in the face of difficulties, one that needs challenge rather than the assurance of success.

The achievement score (developed by McClelland et al., 1953) derived from stories given to pictures of the Thematic Apperception Test (TAT; Murray, 1936, 1938) is generally considered the most useful measure of the strength of the subject's motive to achieve. Therefore we have drawn our measuring techniques for n Ach from those methods. (For the relation between this and other methods see e.g. DeCharms et al., 1955.)

Resultant achievement motivation:

Given our hypothesis and our theoretical model, the expected higher achievement motivation in the experimental class than in the control class would entail a combination of lower anxiety and higher need to achieve.

However, low anxiety, high n Ach students or, as Cronbach (1967) calls them, "constructively motivated" as opposed to the "defensively motivated" high anxiety low n Ach students show a higher school performance (Atkinson and Litwin, 1960) especially in classroom conditions which are less structured and put less emphasis on marks as immediate feedback showing the student whether their work is right or wrong. They also work better under less pressure (Atkinson and Reitman, 1956; for some other studies, see e.g. Kogan and Wallach, 1964; Grimes and Allinsmith, 1961; Amidon and Flanders, 1961; and Flanders, 1965). Hence, it is supposed here that the special educational environment has helped to develop a personality factor which fits this environment better.

INTRINSIC MOTIVATION

Intrinsic and Extrinsic motivation: A Dividing Line Between So-Called "Traditional" and Really "Modern" Ways of Teaching Mathematics.

It is necessary then to take into consideration that the total motivation to perform any achievement task is attributable to two components:

i. achievement - related (or intrinsic) motivation (motivation which is inherent in information processing and action; Hunt, 1971); and

ii. extrinsic motivation (the origin of which is outside of the information processing and action; Feather, 1961).

The traditional way of teaching mathematics emphasizes mainly the latter. One of the main extrinsic motivation is marking, which is placed in the centre of traditional education. This is the chief motivation of pupil's work; they try to get good marks and avoid getting poor ones. Marking outlines the goals of teaching more firmly than the actual working of the syllabus does. As a result, more emphasis is laid on those sections of the subject that call for algorithmic skills than those which need inventiveness; pupils are motivated to learn easily forgettable details rather than to understand connections; more intelligent students, easily getting the highest marks, reach an unjustified satisfaction, while less intelligent students, fearing the lowest mark, become unnecessarily anxious.

Lack of curiosity and internal motivation for learning is often the result of years of teaching which has so emphasized extrinsic incentives that the student has few opportunities for finding intrinsic satisfactions (McKeachie, 1961).

Given the accelerating increases in sheer quantity of information available in this century, it seems all the more important to us to stress internal motivation. The product of traditional, externally motivated schools is often a person who makes little effort to obtain - or even avoids - new knowledge in the fields he studied in school, whereas the encouragement of internal motivation in a student will help him to continually update his learning.

Several efforts in education seem to take these facts into consideration. As Day, Berlyne, and Hunt (1971) write in the preface to their recently published book of papers presented at a symposium on intrinsic motivation in education at the Ontario Institute for Studies in Education:

> Education is moving in the direction of greater concern with motivational variables in the teaching-learning process. There is evidence of a realization that intrinsic motivation is a useful tool in the hands of a skilled teacher. Through its judicial manipulation, he can increase learning while decreasing discontent and the negativism that exists in education today. He can make education interesting and rewarding in itself, so that the student remains excited by the learning process long after he has left the halls of formal education.

The OPI mathematics project has hoped to offer the pupils encouragement by revealing such intrinsic qualities of the subject as the recognition of familiar relationships in new circumstances, the discovery of relationships of a similar nature in seemingly remote fields, and the recognition of the harmony of the created system. Such perceptions can offer a real aesthetic pleasure (Suranyi et al., 1972).

Intrinsic and extrinsic motivation: Which is More Effective?:

It is not easy to answer such a question, first because very little research has been done in this field, and secondly because the question is really quite like "Which is the more powerful, electric motors or steam engines?" The answer obviously depends on how big an electric motor and how big a steam engine is to be considered.

As Berlyne (1971; p. 193) writes:

> There might, however, be some sense in asking whether the most powerful electric motor that has so far been constructed

will out perform the most powerful steam engine in existence, but this does not prejudge the issue at all with regard to motors or engines that have still to be designed. And there is the important fact that electric motors are on the whole more efficient than steam engines. That is to say, they produce the same power at less cost and they make fuller use of the thermal energy on which they draw. The corresponding questions can appropriately be asked with regard to intrinsic motivation. In view of the zest for activity and for inquiry that characterize chidren when they are not in school, the analogy between intrinsic motivation and hydroelectric power has some appeal. They both utilize resources that nature provides in abundant quantities and that would otherwise go to waste. Education based on intrinsic motivation has the further attraction that it seems to be enjoyable. All things being equal, we should certainly prefer children to have a good time in school than to find school distasteful, boring and exhausting.

Intrinsic motivation and the need to achieve:

The relationship between *n* Ach and intrinsic motivation is also not simple. There may even be a conflict between interest in a subject for its own sake and desire for achievement (Vernon, 1969; Car[illegible]on, 1963). But if we assume that extrinsic pushes and pulls in a school situation will have the effect of increasing student anxiety, then we could conclude that the expected effect of the composite mathematics learning method (that is, a parallel increase in *n* Ach and decrease in anxiety) will largely be the result of the students' higher intrinsic motivation.

22.2.2 Method

A modified version (McClelland et al., 1953) of the Thematic Apperception Test (TAT: Murray, 1938).*

As several studies (Ricciuti and Clark, 1954; Reitman and Atkinson, 1958) using the TAT test have shown, there is a less predictive validity of TAT stories late in a series so it is recommended to use only 3 to 6 pictures. Although Atkinson (1958) recommended several sets of pictures, for our special purpose we chose a separate set, consisting of 5 pictures chosen from the TAT (one is different for boys and girls) and 1 specially designed: A blurred photo of the children in a typical Hungarian class. (There is some evidence that pictures related to the special interests of the groups performing the test are more useful to differentiate between n Ach scores. (For school situation, see e.g. Nathan and Mauco, 1966.)

The TAT pictures used in our study were the following:

3BM	Figure crumpled on floor, resting head on divan.
7BM (only for boys)	Young man and old man.
7GF (only for girls)	Older woman sitting on sofa with girl holding a doll
13B	Little boy sitting on the doorstep of a log cabin.
14	Silhouette in opening
17BM	Naked man clinging to a rope.

---

* For a full description of this procedure see e.g. McClelland (1951).

Young children being the subjects, we did not use the generally recommended group method with slide presentation, but carried out the examination individually with the original TAT pictures supplemented with the photo mentioned before. We told the subjects that they had 5 minutes to tell their stories but we did not in fact take this time limit into consideration. To eliminate the effect of the serial position of the pictures we randomized their order of appearance for each subject.

The instructions and the scoring system were the same as those recommended by McClelland, Atkinson et al. (1953). The TAT test was used only in grades 3, 4 and 5.

The CMAS Test, used for measuring anxiety, is described under hypothesis 10.

TREATMENT OF DATA

No mathematical statistical methods were used. Separate graphs were constructed for girls and boys to compare experimental and control n Ach scores.

EXPECTATION

Higher scores in the experimental classes than in the control classes in our TAT-based measurement would show greater need to achieve (n Ach). Lower CMAS scores in the experimental classes would show less avoidance of failure. According to our model, a combination of higher n Ach scores and lower CMAS scores in the experimental classes would show a greater achievement motivation, thus supporting hypothesis 11 in the Budapest study.

As in apparent in Figures 22-1 and 22-2 a higher n Ach score was obtained in experimental classes than in control classes in all grades for both girls and boys. These higher scores were neither a result of longer stories in the experimental classes (the average number of words used in a story ranged between 50 [in picture 14] and 60 [in picture 7] with an average of 55, about the same in experimental and control classes), nor a result of a strikingly high difference in one category, but a result of quite consistently higher scores in each category taken into consideration. So these results may reflect a really greater need for achievement in the experimental classes.

Because significantly lower anxiety scores were found in the experimental class than in the control class in grade 4, we have fairly good evidence that there is greater achievement motivation in the experimental class in this grade. As lower (though not significantly lower) anxiety scores were obtained in the experimental class than in the control class in grade 3, our results support hypothesis 11 in this grade as well. In grade 5 the average anxiety score was slightly higher in the experimental class, but the n Ach scores were so much in favour of the experimental class (especially for the boys) that we can at least state that the results even in this grade do not oppose our hypothesis.

22.3 Effects on Personality

Hypothesis 12: Learning mathematics with the Dienes method causes a favourable change in the student's personality.*

---

* This variable was examined only in the Budapest Study.

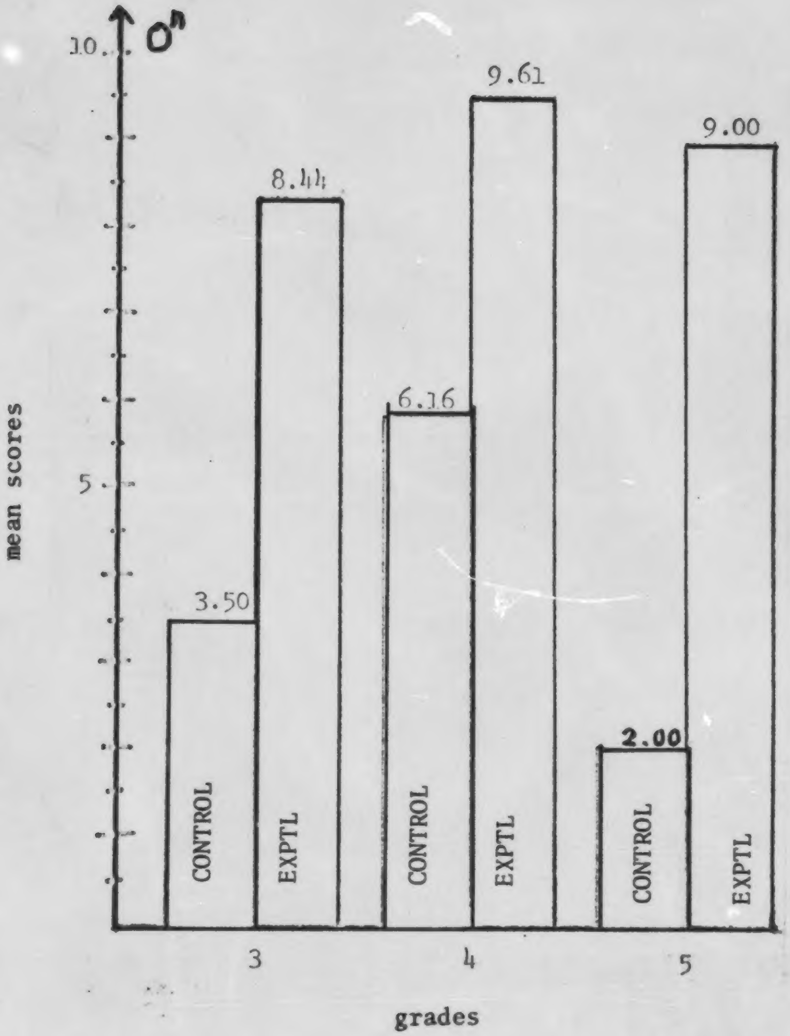


Fig. 22-2. n Ach scores for boys.

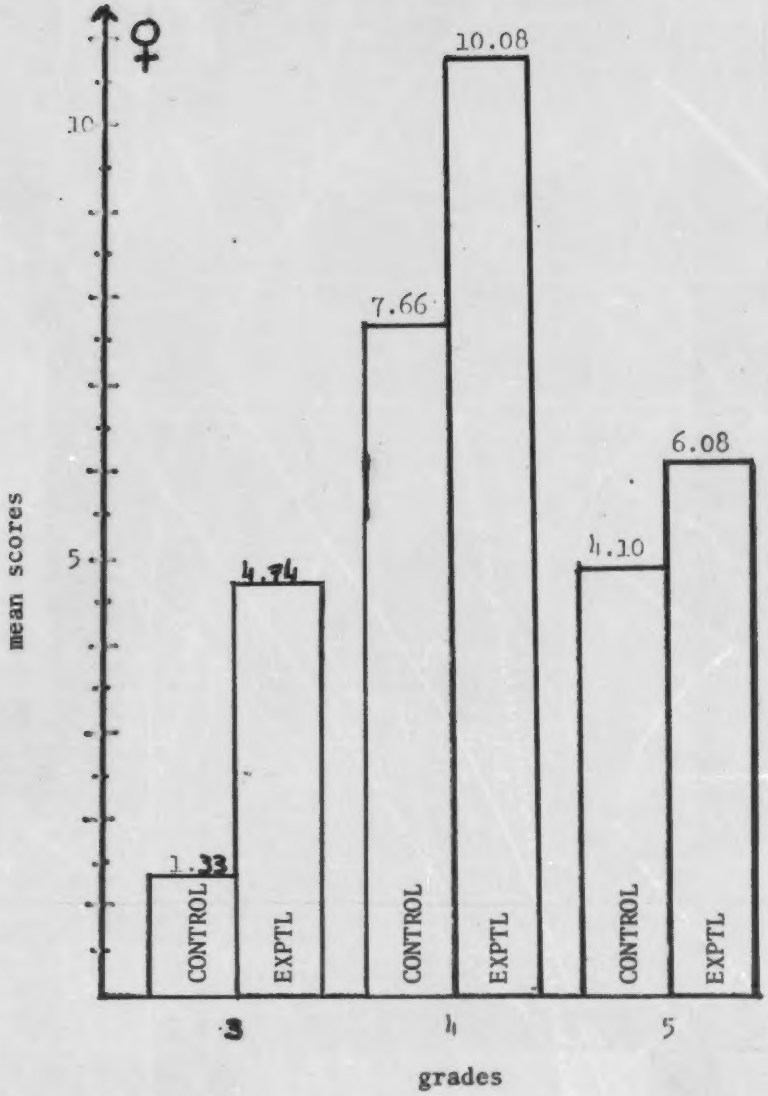


Fig. 22-3. n Ach scores for girls.

22.3.1 Theoretical background

It is especially difficult to account for the inclusion of a personality test in our evaluation. As Glass has pointed out very sharply in a discussion about criterion variables in evaluation:

"There is no evidence for an interaction of curriculum treatments and personalogical variables." (Wittrock and Wiley, 1970; p. 210.)

The results of the studies carried out in this field are generally ambiguous, leaving open the question whether the differences found in the student's personality originate from the educational treatment, or from the general human characteristics. One reason for this could be that both educational treatment and human characteristics are very complex and broadly defined variables. To get more definite results the researcher should break these down into very small specific variables.

After Glass an example of such specific variables would be the tempo of presentation of a programmed instruction on the one side and the students characteristic working tempo on programmed instruction on the other (study by Kress and Gropper, 1966).

But also for more complex personality variables, there is some evidence that they play a role in the effectiveness of instruction; at least as so called moderator variables, that means that students possessing a trait in a higher degree may be affected more by a certain teaching method than students with a lower degree. Usually effectiveness has been determined in terms of achievement scores only, not on the personality trait itself, which takes the edge off Glass' statement. In the following passages some studies are reported which are concerned with personality traits in the field of school mathematics (after Glennon and Callahan, 1968).

> Hebron (1962), in examining his data on relationships between learning a new number system and certain personality dimensions, suggested that "extrovert" attitudes in learning may favor the assimilation of the first elementary facts of a new situation, while "introverts" may be more capable, when this stage is passed, in applying these facts in more complex problems. He goes on to suggest that facility in the new quantitative field is most enjoyed by the extrovert, while the introvert is more at ease operating with familiar symbols of a verbal nature.

> In studying the dogmatic person and his critical thinking ability, Kemp (1960) found the "high dogmatics" have the greater percentage of errors in problems which required the studying of several factors or criteria for decision and the deferring of a conclusion until each factor has been judiciously considered.. He writes that apparently the "high dogmatic" has difficulty in tolerating ambiguity and is thus impelled toward a "closure" before full consideration is given to each piece of contributing evidence.

A relationship between self concept and achievement in elementary school mathematics could also be shown:

> The results of a study by Fink (1962) appear to confirm the hypothesis that a relationship does exist between adequacy of self-concept and level of academic achievement. This conclusion appears to be unquestionable for boys, considerably less so for girls.

> Bodwin (1957), in a study using third- and sixth-grade students, found a positive and significant relationship (.78 on the third-grade level and .68 on the sixth-grade level) between immature self-concept and arithmetic disability.

However, a relationship between self-concepts and mathematics learning as pointed out by a correlation coefficient does not tell us where the cause is lying on the one side or the other, or underlying both. In the frame of evaluation, there is a study by Wilson, Stuckey, and Langevin (1972) which could demonstrate that the children's self-concept (attitude towards the self) was more favourable in an open plan school than in a traditional one. A number of studies are concerned with cognitive styles as moderator variables and, through their findings, point out the need for a more individualized instruction for learning.

> Kagan *et al.* (1966) have been concerned with conceptual tempo of children. They have found that impulsive children will tend to report the first hypothesis that occurs to them, and this response is often incorrect. The reflective child, on the other hand, delays a relatively long time before reporting a solution and is usually correct.
>
> Programs that emphasize the discovery method of instruction insist that practice in inference is the key aim of the educational enterprise. Many teachers believe an incorrect inference shows insufficient knowledge and usually do not appreciate the role of an impulsive attitude in determining the quality of the inferential process. Kagan and his colleagues suggest that it may be profitable to consider training the child in reflection when facts and rules are introduced in order to facilitate the general quality of a student's performance.

While being trained in a certain way of thinking could be one of the possible side effects of the Dienes method of mathematics learning and therefore deserving of study, lacking the appropriate tests we directed our attention to broader personality traits, as can be easily assessed by Cattell's Personality Questionnaire. The reason for the assumption that differences could be found on such dimensions was given by research studies such as the above. Though the possibility of getting significant results is not very high because of the complexity of the variables involved, the assessment of personality traits was included in order to allow for a differential analysis of some of the other data by using this assessment moderator variable.

22.3.2 Method

To assess students personality traits, Cattell's Children's Personality Questionnaire (CPQ, 1963 edition, French translation 1968) was used.

Generally the CPQ yields a general assessment of personality by measuring fourteen distinct dimensions or traits of personality which have been found by psychologists to approach the total personality (Cattell, 1946, Cattell and Warburton 1967). These dimensions are functionally independent as established by factor-analytic research. By working with these fourteen scores, individually and in combination, it is possible to obtain among others predictions of school-related criteria such as academic achievement, especially under-achievement.

The dimensions are described in Table 22-6. Each dimension is defined by two poles or extremes. The left side (below) gives descriptions for a lower score on the factor or dimension, while the right side describes what the high-scoring child tends to be like.

The CPQ was administered to children of all 3 grades as a group test. Both parts (second part 5-8 days later) of Form A were used. Each part consists of 140 questions, for example, question 37 asks:

"Would you rather collect stamps ☐ or ☐ play football?"

Comparison was made on the raw scores of each factor which may vary between 0 and 10.

22.3.3 Expectation

We expected significant differences between experimental and control classes on most of the personality traits in the CPQ, as side effects of the Dienes method. We expected higher average scores in the experimental classes in Factors A, B, C, D, E, F, H, while we expected higher average scores in the control classes in Factors G, J, N, O and Q_4.

22.2.4 Results

The results are shown in Tables 22-7 and 22-8. As can be seen, the results on the CPQ did show some interesting differences between girls and boys: Girls were more obedient, mild, conforming, submissive (Factor E), more tender-minded, dependent, overprotected, sensitive (Factor I), and more forthright, natural, artless, sentimental (Factor N), while boys were slightly more excitable, impatient, demanding, overactive (Factor D).

The CPQ failed to show the significant experimental-control differences we expected (see Table 22-8). Even the few significant results are quite controversial: In Factor A the grade 4 experimental class proved to be more warmhearted, outgoing, participating, easygoing which means, according Cattell, that "they respond more favourably to their teachers and to the school situation generally". On the other hand, on the same factor the grade 5 control class revealed these traits to a lower degree - which underlines again the importance of the teacher. (See also page 199 on the students' attitude.)

Contrary to what we expected, the children in the control classes proved to be "more intelligent" (Factor B). It is, however, a very unreliable measurement. As Cattell himself writes: "It is not intended to replace the more valid estimate of IQ obtained from longer measures such as the Culture Fair Intelligence Test." It should be pointed out that this factor emphasizes concrete versus abstract thinking and convergent versus divergent thinking.

We have no real explanation at present for the results which show that grade 3 experimental children are more affected by feelings, emotionally less stable, easily upset, changeable, of lower ego strength and therefore more lacking in frustration tolerance than grade 3 control children, and that the latter showed themselves to be more assertive, independent, aggressive, stubborn and dominant. Nor can we explain the fact that grade 5 control children are emotionally stable, ready to face reality, calm, of higher ego strength and therefore socially more mature and better prepared to cope effectively with others than grade 5 children in the experimental group (Factor C).

TABLE 22-6
Personality Factors from Cattell's (1968) Children Personality Questionnaire, Manual.

Factor	DESCRIPTION OF FACTOR	
	Low Score	High Score
A	Reserved, Detached, Critical, Cool	Warmhearted, Outgoing Participating, Easygoing
B	Less Intelligent, Concrete-thinking, Lower Scholastic Mental Capacity	More Intelligent, Abstract-thinking, Bright, Higher Scholastic Mental Capacity
C	Affected by Feelings. Emotionally Less Stable, Easily Upset, Changeable, Lower Ego Strength	Emotionally Stable, Faces Reality, Calm, Higher Ego Strength
D	Phlegmatic, Deliberate, Inactive, Stodgy	Excitable, Impatient, Demanding, Overactive
E	Obedient, Mild, Conforming, Submissive	Assertive, Independent, Aggressive, Stubborn Dominant
F	Sober, Prudent, Serious, Taciturn	Happy-go-lucky. Impulsively Lively, Gay, Enthusiastic, Heedless
G	Expedient. Disregards Rules, Undependable, Bypasses Obligations, Weaker Superego Strength	Conscientious, Persevering, Staid, Rule-bound, Stronger Superego Strength
H	Shy. Restrained, Diffident Timid	Venturesome, Socially Bold, Uninhibited, Spontaneous
I	Tough-minded. Self-reliant, Realistic, "No-nonsense"	Tender-minded. Dependent, Overprotected, Sensitive
J	Vigorous. Goes Readily with Group, Zestful, Given to Willing Action	Circumspect. Obstructive, Individualistic, Unwilling to Act with Group
N	Forthright. Natural, Artless, Sentimental	Shrewd. Calculating, Astute
O	Self-assured. Placid, Secure, Serene, Untroubled Adequacy	Apprehensive, Worrying. Depressive, Troubled, Insecure
O_3	Casual. Careless of Social Rules, Untidy, Follows own Urges, Low Integration	Controlled, Socially-precise, Self-disciplined, Compulsive, High Self-concept Control
O_4	Relaxed. Tranquil, Torpid, Unfrustrated	Tense. Driven, Overwrought Fretful

Table 22-7

The mean-differences between boys and girls on the 14 personality factors as assessed by Cattell's "Children's Personality Questionnaire (C P Q)".

See footnotes at the end of Table 22-8.

Groups		Boys			Girls						
Factors	Statistics Grades	$\overline{X}$	SD	N	$\overline{X}$	SD	N	$t_{crit}^{.05}$	t	$F_{crit}^{.10}$	F
A	5	7.44	1.34	18	7.00	1.54	22	2.02	0.98	2.23	1.31
	4	6.22	2.10	23	7.46	1.39	13	2.12	-1.85	2.05	2.28
	3	7.22	1.81	18	7.53	1.39	19	2.02	-0.57	2.25	1.70
	Σ^1	-0.15	1.83	59	0.16	1.50	54	1.98	-0.99	1.58	1.49
B	5	6.83	1.50	18	6.64	2.03	22	2.02	0.35	2.23	1.84
	4	5.39	2.08	23	5.46	2.10	13	2.03	-0.10	2.20	1.02
	3	4.94	1.81	18	4.95	2.52	19	2.02	-0.09	2.25	1.94
	Σ	0.02	1.84	59	-0.02	2.23	54	1.98	0.12	1.56	1.48
C	5	6.94	1.58	18	6.14	1.71	22	2.02	1.55	2.23	1.18
	4	6.09	1.95	23	7.38	1.78	13	2.03	-2.03^{*}	2.05	1.21
	3	6.61	1.89	18	6.95	2.04	19	2.02	-0.52	2.25	1.16
	Σ	-0.10	1.87	59	0.11	1.91	54	1.98	-0.59	1.56	1.04
D	5	3.67	1.60	18	3.05	1.43	22	2.02	1.28	2.15	1.25
	4	3.91	1.95	23	2.62	1.21	13	2.12	2.11	2.05	2.60
	3	3.56	1.83	18	3.11	1.65	19	2.03	0.78	2.25	1.23
	Σ	0.36	1.82	59	-0.39	1.49	54	1.98	2.40^{*}	1.58	1.50
E	5	5.28	1.19	18	2.68	1.22	22	2.02	6.78^{***}	2.23	1.05
	4	4.17	1.52	23	2.92	1.69	13	2.03	2.21^{*}	2.20	1.23
	3	3.89	1.56	18	3.21	1.51	19	2.02	1.34	2.25	1.07
	Σ	0.72	1.52	59	-0.78	1.49	54	1.98	5.31^{***}	1.58	1.04
F	5	5.33	1.49	18	2.41	1.30	22	2.02	6.53^{***}	2.15	1.31
	4	4.96	1.78	23	1.85	1.03	13	2.12	5.62^{***}	2.05	3.01
	3	5.06	1.96	18	2.58	1.90	19	2.02	3.90^{***}	2.25	1.06
	Σ	1.32	1.77	59	-1.44	1.52	54	1.98	8.90^{***}	1.58	1.35
G	5	7.06	1.22	18	7.23	1.20	22	2.02	-0.44	2.15	1.03
	4	6.91	2.38	23	7.62	1.64	13	2.12	-0.92	2.05	2.09
	3	7.11	1.79	18	7.21	1.47	19	2.02	-0.18	2.25	1.48
	Σ	-0.14	1.91	59	0.16	1.42	54	2.00	-0.93	1.58	1.80

Table 22-7 cont.

The mean-differences between boys and girls on the 14 personality factors as assessed by Cattells' "Children's Personality Questionnaire (C P Q)".

Groups		Boys			Girls						
Factors	Statistics / Grades	$\overline{X}$	SD	N	$\overline{X}$	SD	N	$t_{crit}^{.05}$	t	$F_{crit}^{.10}$	F
H	5	6.28	1.66	18	5.18	1.92	22	2.02	1.93	2.23	1.34
	4	5.74	1.87	23	5.62	1.60	13	2.03	0.21	2.05	1.37
	3	6.11	1.88	18	5.74	2.15	19	2.02	0.56	2.25	1.30
	Σ	0.26	1.83	59	-0.28	1.94	54	1.98	1.53	1.56	1.13
I	5	3.94	1.61	18	6.32	1.46	22	2.02	-4.83***	2.15	1.23
	4	3.57	1.50	23	5.85	1.70	13	2.03	-4.03***	2.20	1.29
	3	4.39	1.80	18	5.79	2.33	19	2.02	-2.05*	2.25	1.68
	Σ	-0.94	1.65	59	1.03	1.89	54	1.98	-5.87***	1.56	1.31
J	5	3.67	1.37	18	4.55	1.78	22	2.02	-1.76	2.23	1.67
	4	4.09	1.77	23	4.62	1.39	13	2.03	-0.99	2.05	1.62
	3	3.89	1.15	18	3.32	1.66	19	2.02	1.23	2.25	2.08
	Σ	-0.13	1.51	59	0.14	1.68	54	1.98	-0.92	1.56	1.23
N	5	4.44	1.74	18	3.27	1.39	22	2.02	2.32*	2.15	1.57
	4	3.39	1.81	23	2.77	1.62	13	2.03	1.06	2.05	1.24
	3	4.00	1.41	18	2.58	1.73	19	2.02	2.75**	2.25	1.49
	Σ	0.51	1.69	59	-0.55	1.58	54	1.98	3.45***	1.58	1.15
O	5	3.22	1.40	18	4.36	1.97	22	2.02	-2.14*	2.23	1.98
	4	3.52	1.77	23	3.54	2.17	13	2.03	-0.02	2.20	1.51
	3	3.06	2.07	18	3.05	1.99	19	2.02	0.00	2.25	1.08
	Σ	-0.19	1.79	59	0.21	2.04	54	1.98	-1.12	1.56	1.30
Q_3	5	6.17	1.83	18	6.09	1.35	22	2.02	0.15	2.15	1.86
	4	5.43	2.53	23	5.92	2.64	13	2.03	-0.54	2.20	1.09
	3	5.67	1.53	18	6.37	1.81	19	2.02	-1.28	2.25	1.41
	Σ	-0.17	2.07	59	0.18	1.90	54	1.98	-0.93	1.58	1.18
Q_4	5	4.67	2.08	18	4.27	1.81	22	2.02	0.63	2.15	1.32
	4	4.22	2.48	23	3.62	1.94	13	2.03	0.81	2.05	1.63
	3	3.67	1.67	18	3.74	1.58	19	2.02	-0.13	2.25	1.11
	Σ	-.14	2.14	59	-0.15	1.78	54	1.98	0.79	1.58	1.45

Table 22-8 The mean-differences between experimental and control groups on the 14 personality factors as assessed by Cattells "Children's Personality Questionnaire (C P Q)".

GROUPS Statistics		EXP			CONTROL			EXP + CONTL						
Personality Factors	Grades	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$t^{.05}_{crit}$	t	$F^{.10}_{crit}$	F
A reserved, warmhearted	5	6.63	1.09	19	7.71	1.58	21	7.20	1.47	40	2.02	-2.55*	2.16	2.1
	4	7.42	1.27	19	5.82	2.26	17	6.67	1.97	36	2.09	2.58*	2.30	3.1
	3	7.44	1.67	18	7.32	1.56	19	7.38	1.62	37	2.02	0.24	2.25	1.1
	Σ[1]	0.08	1.46	56	-0.08	1.88	57	0.00	1.69	113	2.00	0.52	1.56	1.6
B less intelligent/more intelligent	5	6.26	1.71	19	7.14	1.81	21	6.72	1.82	40	2.02	-1.58	2.19	1.1
	4	4.79	1.85	19	6.12	2.11	17	5.42	2.09	36	2.03	-2.00	2.25	1.3
	3	4.44	2.27	18	5.42	2.03	19	4.95	2.20	37	2.02	-1.38	2.25	1.2
	Σ	-0.53	1.95	56	0.52	1.98	57	0.00	2.04	113	1.98	-2.84**	1.56	1.0
C affected by feelings emotionally stable	5	7.21	1.73	19	5.86	1.39	21	6.50	1.70	40	2.02	2.70*	2.16	1.5
	4	6.89	1.62	19	6.18	2.28	17	6.56	1.99	36	2.03	1.08	2.19	1.9
	3	6.06	2.07	18	7.47	1.60	19	6.78	1.97	37	2.02	-2.32*	2.25	1.6
	Σ	0.12	1.91	56	-0.12	1.86	57	0.00	1.89	113	1.98	0.68	1.56	1.0
D phlegmatic - excitable	5	3.00	1.69	19	3.62	1.33	21	3.32	1.54	40	2.02	-1.28	2.16	1.6
	4	3.68	1.92	19	3.18	1.69	17	3.44	1.83	36	2.03	0.84	2.30	1.2
	3	3.00	1.56	18	3.63	1.87	19	3.32	1.76	37	2.02	-1.12	2.25	1.4
	Σ	-0.13	1.75	56	0.13	1.65	57	0.00	1.71	113	1.98	-0.82	1.56	1.1
E obedient - assertive	5	4.00	1.49	19	3.71	1.98	21	3.85	1.77	40	2.02	0.52	2.19	1.7
	4	3.89	1.68	19	3.53	1.68	17	3.72	1.69	36	2.03	0.65	2.25	1.0
	3	2.83	1.42	18	4.21	1.40	19	3.54	1.57	37	2.02	-2.97*	2.25	1.0
	Σ	-0.12	1 59	56	0.12	1.76	57	0.00	1.68	113	1.98	-0.74	1.56	1.2
F sober - happy-go-lucky	5	3.68	1.84	19	3.76	2.16	21	3.72	2.01	40	2.02	-0.12	2.19	1.8
	4	3.79	2.07	19	3.88	2.25	17	3.83	2.15	36	2.03	-0.13	2.25	1.1
	3	3.83	2.79	18	3.74	1.68	19	3.78	2.29	37	2.06	0.12	2.25	2.7
	Σ	-0.01	2.26	56	0.01	2.04	57	0.00	2.15	113	1.98	-0.06	1.56	1.2
G expedient - conscientious	5	7.16	1.35	19	7.14	1.08	21	7.15	1.22	40	2.02	0.04	2.16	1.5
	4	7.26	1.68	19	7.06	2.60	17	7.17	2.17	36	2.09	0.27	2.25	2.4
	3	6.94	1.65	18	7.37	1.60	19	7.16	1.64	37	2.02	-0.79	2.25	1.1
	Σ	-0.03	1.57	56	0.03	1.82	57	0.00	1.70	113	1.98	-0.21	1.56	1.3

Table 22-8 cont.

GROUPS / Personality Factors	Statistics / Grade	EXP X	EXP SD	EXP N	CONTROL X	CONTROL SD	CONTROL N	EXP + CONTL X	EXP + CONTL SD	EXP + CONTL N	$t_{crit}^{.05}$	t	$F_{crit}^{.10}$	F
H shy/ venturesome	5	5.32	1.72	19	6.00	1.98	21	5.67	1.89	40	2.02	-1.17	2.19	1.3
	4	5.58	1.63	19	5.82	1.92	17	5.69	1.78	36	2.03	-0.41	2.25	1.3
	3	6.00	2.08	18	5.84	1.98	19	5.92	2.03	37	2.02	0.24	2.25	1.1
	Σ	-0.13	1.83	56	0.13	1.97	57	0.00	1.90	113	1.98	-0.75	1.56	1.1
I tough-minded/ tender-minded	5	5.16	2.01	19	5.33	1.86	21	5.25	1.93	40	2.02	-0.29	2.16	1.1
	4	4.37	1.56	19	4.41	2.25	17	4.39	1.92	36	2.03	-0.07	2.23	2.0
	3	5.28	2.33	18	4.95	2.06	19	5.11	2.20	37	2.02	0.46	2.25	1.2
	Σ	0.02	1.99	56	-0.02	2.05	57	0.00	2.02	113	1.98	0.09	1.56	1.0
J vigorous/ circumspect	5	4.16	2.23	19	4.14	0,89	21	4.15	1.67	40	2.08	0.03	2.16	6.3
	4	4.74	1.48	19	3.76	1.70	17	4.28	1.66	36	2.03	1.82	2.25	1.3
	3	3.78	1.69	18	3.42	1.18	19	3.59	1.46	37	2.02	0.74	2.25	2.0
	Σ	0.22	1.84	56	-0.21	1.29	57	0.00	1.60	113	2.00	1.43	1.56	2.0
N forthright/ shrewd	5	3.89	1.48	19	3.71	1.80	21	3.80	1.66	40	2.02	0.35	2.19	1.4
	4	3.00	1.59	19	3.35	1.94	17	3.17	1.77	36	2.03	-0.59	2.25	1.4
	3	3.39	1.77	18	3.16	1.69	19	3.27	1.73	37	2.02	0.41	2.25	1.0
	Σ	0.01	1.62	56	-0.01	1.81	57	0.00	1.72	113	1.98	0.08	1.56	1.2
O self-assured/ apprehensive	5	3.47	1.90	19	4.19	1.68	21	3.85	1.82	40	2.02	-1.26	2.16	1.2
	4	3.53	2.01	19	3.53	1.82	17	3.53	1.92	36	2.03	-0.00	2.30	1.2
	3	3.06	1.90	18	3.05	2.14	19	3.05	2.03	37	2.02	0.00	2.25	1.2
	Σ	-0.13	1.95	56	0.13	1.89	57	0.00	1.92	113	1.98	-0.70	1.56	1.0
Q_3 casual/ controlled	5	6.47	1.60	19	5.81	1.50	21	6.12	1.58	40	2.02	1.35	2.16	1.1
	4	4.95	2.21	19	6.35	2.76	17	5.61	2.58	36	2.03	-1.67	2.25	1.5
	3	6.28	1.69	18	5.79	1.70	19	6.03	1.72	37	2.02	0.87	2.25	1.0
	Σ	-0.03	1.91	56	0.03	2.07	57	0.00	2.00	113	1.98	-0.14	1.56	1.1
Q_4 relaxed/ tense	5	4.11	1.83	19	4.76	2.00	21	4.45	1.95	40	2.02	-1.08	2.19	1.1
	4	4.21	2.33	19	3.76	2.29	17	4.00	2.32	36	2.03	0.58	2.30	1.0
	3	3.72	1.79	18	3.68	1.45	19	3.70	1.63	37	2.02	0.07	2.25	1.5
	Σ	-0.04	2.02	56	0.04	1.95	57	0.00	1.98	113	1.98	-0.21	1.56	1.0

Footnote: The formula for t for heterogenous variances was applied if variances were significant at the 10 percent level ($F > F_{crit}^{.10}$), underlined.

* t significant at the 5 percent level ** t significant at the 1 percent level *** t significant at the 0.1 percent level

1 the method used to combine scores is described in the footnote of Table 15-3.

Chapter 23

Attitude Toward Problems

Hypothesis 13: The Dienes method helps student to face problems rather than to avoid them.

Several new educational methods assume that one of their objectives is to develop in their pupils an attitude towards cognitive problems they meet which is more favourable than mere avoidance. R. Davis (1964), in his discussion of the "discovery method" used in the Madison Project, explains how "traditional education" can fail:

"Ask 5th graders what they can write in the "$\square$" to produce a true statement from the open sentence

$$\square \times \square = 2$$

Then try 1; it is too small: $1 \times 1 = 1 < 2$

Then try 2; it is too large: $2 \times 2 = 4 > 2$

Then try 1.5; it is too large: $1.5 \times 1.5 = 2.25 > 2$

Then try 1.4; it is too small: $1.4 \times 1.4 = 1.96 < 2$

Now they are off to the races, and they produce a great spate of suggestions, approximations, and relevant questions.

Propose the same question to older children; the older the child, the less creative and enthusiastic the response, until finally by the college freshman year, students very frequently respond by saying: I don't think we had that in our high school. That is to say: IF NOBODY EVER TOLD ME, THEN I CANNOT POSSIBLY KNOW.

Problem solving plays an especially important role in mathematics learning. With traditional mathematics lessons, students encounter notorious difficulty especially in solving problems presented in verbal form (Loftus, 1970). This difficulty has long been a worry to mathematics teachers as well as the cause of much pupil anxiety (Riedesel, 1964). Most new mathematics programmes, such as the Dienes method, place much emphasis on developing the students' ability to solve problems. What is perhaps more important, by providing a large variety of interesting problems day after day, these programmes attempt to develop in the children favorable attitudes toward the problems. (For more details, see Klein, 1973a.)

Not having an appropriate measuring device, we constructed a new test, the Sherbrooke Problem Facing Questionnaire (SPFQ) which was expected to be a measure of how well educational methods reach the objective of creating an attitude towards cognitive problems which is more favourable than mere avoidance. Because of this we shall describe the method used to examine hypothesis 13, in greater detail.

23.1 Theoretical background

23.1.1 Problem solving ability

As we mentioned before one of the thorniest problems besetting elementary school teachers is the improvement of verbal problem-solving ability. Indeed, many elementary teachers enjoy less success with this phase of mathematics, science, etc... instruction than with any other phase (Ashlock and Herman, 1970, p. 193).

This is a pity because problem solving is one of the basic activities used in many school subjects, particularly in mathematics. As Rosenbloom (1966, p. 130) says: "Other mathematical activities such as generalization, abstraction, theory building, and concept formation are based on problem solving," and further as P.S. Jones (1966, p.112) writes: "Some of the intangible goals stimulating interest, intuition, creativity, problem solving facility, and understanding of the model making process can be achieved through the improved use of problems." Many prominent educators not only regard problem solving as a basic activity in mathematics, but believe that it should and could be developed (Polya, 1966, Lynes; 1967) with the proper methods.

Problem solving is a widely investigated phenomenon in psychology (see e.g. Duncan, 1959, Kleinmutz, 1966).

Several studies have been conducted on how different teaching methods effect problem solving ability (Wilson, 1967; Lerch and Hamilton, 1966; Lyda and Church (1964), how problem solving ability relates to aptitudes (Anderson 1967); abilities (e.g. ability to converse about numerousness, Steffe, 1968), creativity (Guilford, 1966; 1967b); other intellectual factors (Merrifield et al. 1960; 1962); prior learning (Scandura, 1966a, b); etc. Recently, Guilford (1967a), Guilford-Hoepfner (1971) presented a problem-solving model built on his structure-of-intellect (SI) model (Guilford, 1959) together with several tests designed to measure problem-solving ability.

However, these researchers have only just begun to analyse the results of their studies. The editors of a fairly recent book, Current Research in Instruction (Anderson et al., 1969), expressed a common attitude that is fairly typical among educators towards these psychologists and their evaluations; they have entirely rejected the research done in the problem solving area (and in creativity as well) although the topic is quite popular among educators today, because "In addition to being poor research on the whole, studies dealing with problem solving and creativity have not clarified the features of instruction or conditions of learning, which lead to the desired outcomes."

One reason why the existing studies are greeted with such low regard may be that principles of problem solving as formulated on the basis of laboratory studies are not necessarily co-extensive with or even analogous to principles of meaningful reception learning in mastering subject-matter material in school situations (Ausubel, 1967).

23.1.2 Attitude towards problems

Problem solving ability and attitude towards problems are two very closely related phenomena. A student's attitude towards problems is in itself a very important factor. If we believe that problem-solving should pervade all mathematics curricula (and probably several other subjects as well), then it is basic to realize that in doing so we will not frustrate the children, but help them to develop a more favourable relationship between themselves and the school in which they are learning.

Problems have many values and uses: Problems of interest to the student can motivate his study; problems can test the understanding of concepts, problems can stimulate the student's curiosity; problems can clarify the relationship between what is learned in school and in the world outside the school building, but only if the student during his hours in school learns to like these problems, not to avoid them (see Fig. 23-1).

If we can develop a favourable attitude towards problems in children, then they will be willing take an active role in problem-solving (provided we create a school environment which offers challenging problems), and hence they are

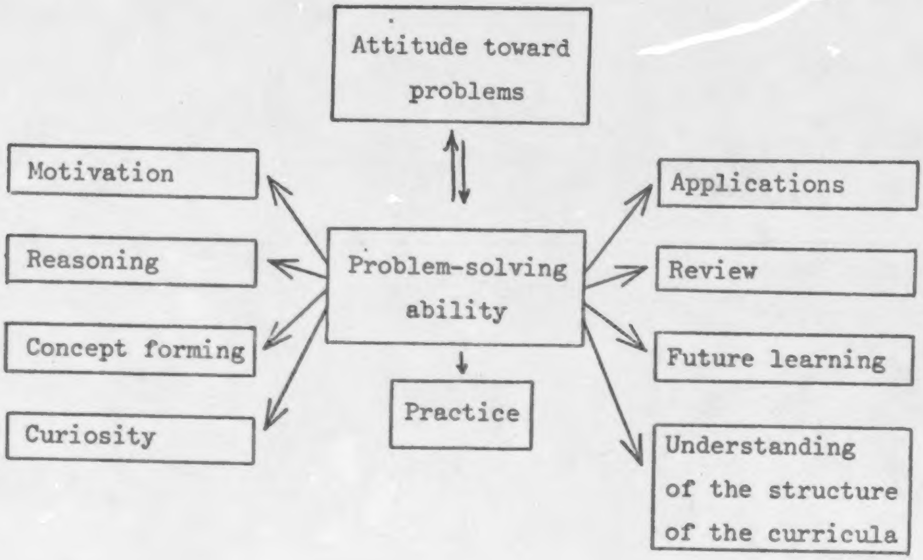

Attitude toward problems
Motivation
Reasoning
Concept forming
Curiosity
Problem-solving ability
Practice
Applications
Review
Future learning
Understanding of the structure of the curricula

Fig. 23-1

The role of problem solving ability and attitude toward problems in curriculum and instruction (after the report of the K-13 Arithmetic-Algebra Committee, 1971).

more likely to develop the power to think logically, to question intelligently, and to apply "the problem solving approach" to problems outside the school, in the "real world" (K-13 Mathematics, 1970).

Psychological literature in the special field of measuring children's attitudes towards problems is rare. In fact, we can cite only one interesting effort in this area: The Child Attitude Inventory for Problem Solving (CAPS), developed by R.S. Crutchfield and M.L. Covington.* This test which is designed to measure the fifth and sixth grade children's attitude towards thinking and problem solving is a group administered paper-and-pencil inventory. It consists of a problem ("How can a flexible and heavy television cable run through a pipe?"), and two scales.

Scale I assesses children's beliefs and attitudes about problem solving. It consists of true-false items measuring variables like such as a child's attitude towards expressing his own ideas, the child's conception of the innateness of problem-solving ability, the wisdom of persisting at difficult problems on which others have failed.

Here are two examples from the statements in Scale I:

1. A problem like the one about the TV cable and the pipe is probably too hard for anyone in the 5th grade to solve.
2. There is probably only one answer to a problem like this one.

---

* According to Johnson and Bommarito (1972) this measure may be obtained from R.S. Crutchfield, Dept. of Psychology, University of California, Berkeley, California, U.S.A. 94700.

In Scale II there are thirty true-false items treating the child's feelings about his own problem solving ability.

Here are two examples form the questions in Scale II:

1. Do you feel that other children in class know more about what to do in working on a problem like this than you do?

2. Would you like to work on a problem like this one?

Questions regarding sources of anxiety about thinking are included in both scales. According to Johnson and Bommarito (1971), some of the statistical characteristics are:

1. The reliability coefficients of CAPS measured by the test-re-test method (with a repeated administration five weeks apart) were around 0.7.

2. On the basis of item analyses on an earlier form of CAPS the revised form of the scales includes highly differentiating items (cf. p. 213).

3. A moderate correlation between the two subscales ($r = .35$ for 325 subjects) shows that the amount of overlap is small and that the two tests measure different variables.

4. The scale scores show a significant and positive correlation with the California Mental Maturity Scale ($r = .33$, $p < .01$).

5. Low negative correlations exist between these scales and two anxiety measures: the Test Anxiety Scale for Children (Mandler and Sarason, 1952) and General Anxiety Scale for Children (Castaneda et al., 1956a, Castaneda, 1961).

Some successful attempts were made to adopt CAPS to college level*.

---

* Personal communication from Professor Jeremy Kilpatrick, Teachers College, Columbia University.

However, because this test consists of direct questioning of the subjects, asking about their behaviour in certain situations, we felt that it is not adequate for very young children. This need for a more appropriate measuring instrument influenced us to develop our problem-facing questionnaire.

23.2

Method

SHERBROOKE PROBLEM FACING QUESTIONNAIRE

SPFQ: French Version, Klein (1973b).

Substantially, the same method was used in the Sherbrooke study and in the New York Study. The method is described in detail here, leaving for the next two sections only those details which differ in the two studies.

23.2.1 The Test Development Process

As a means of developing the original idea, a brain storming session took place in April 1972 at the Centre de Recherches en Psycho-Mathématique, Université de Sherbrooke, Québec, Canada.* The problems proposed had to be such that:

(i) they were intelligible to a child between 9 and 12.
(ii) they could not be solved with techniques the child already knew, and
(iii) the child could find a reasonable, rough estimate of the answer.

After subjective consideration the problems considered best to meet these requirements were arranged temporarily in two parallel forms (Form A and Form B) based solely on psychological judgement (Gulliksen, 1967, p. 174). [To items in Form A and Form B together we shall in what follows refer to as Form (A + B).]

---

* The author is grateful to all members of the Centre whose help made it possible to construct this fairly new test in the relatively short time available to us.

A preliminary scoring system was worked out and after a pre-trial the test was carried out with 108 fourth and fifth grade students as a part of the Sherbrooke study.

The testing was repeated half a year later to get more information concerning the reliability of the test (see 23.2.5). The scoring system was improved (see 23.2.4), and after an item analysis (see 23.2.5) the remaining items from the provisional Form (A + B) were rearranged in a new form. This new form was called form S when based on a mixed grade 4 and 5 population, Form S_4 when based only on a grade 4 and Form S_5 when based only on a grade 5 population.

The American Version of the test was developed during a study in Cedar Falls, Iowa, U.S.A. This study resulted in form R_5. The test was also used in the New York Study.

23.2.2 Description of the Test

The Sherbrooke Problem Facing Questionnaire (SPFQ) consists of two parts. (It is possible to use Part I only without Part II, although Part II provides some additional information.)

PART I

Part I consists of questions and categories for each in which the answers to each question can be given.

Questions: Form A and Form B of the test consist of 22 questions each. Consequently, Form (A+B) consists of 44 questions. Using the item analysis described in 12.2.5.3. We selected 27 questions for Form S_4 and 27 questions for Form S_5 from Form (A+B).

There are two different kinds of questions in the test, Control Questions and Problem Questions.

a) Control Questions

These are fairly easy questions for which a definite or exact answer could be given such as in Example 1 on the front page of the test:

"If one chocolate bar costs 10 cents, how much do three chocolate bars cost?"

There are only a few questions like this in the test: 3 in Form A, in Form B, and in form S; and consequently 6 in form (A + B). These questions do not count in scoring, and their only role is to check whether the subject has the minimal intelligence to participate in the testing and is able to understand the instructions. (If he does not give answers in the appropriate places in at least two out of three questions, then his questionnaire cannot be scored.)

b) Problem questions

The remaining 22 questions have no exact answer. However, it is possible to determine some limits between which the answer falls, as, for example, in Problem 2, on the front page of the test:

> "John bought some sugar. He doesn't remember exactly how much sugar he bought, but he knows that he bought more then two pounds and less than four pounds. Sugar costs 10 cents a pound. How much did John pay for the sugar he bought?"

It is also possible to give an approximation, as in Problem 3:

> "We have three bottles of milk and in each bottle there is about one quart of milk. Perhaps there is a bit less than a quart in each bottle, or perhaps a bit more, but in each bottle there is about one quart. How much milk is there all together in the three bottles?"

Sometimes the question itself gives enough information to make this approximation or gives the limits between which the answer must lie; sometimes again the answer can only be based on a common sense judgement:

> "How much food can a lion eat in a month?"

The subject gives his answers in five columns. He has to place an X in the first column if his answer is "I haven't learned that", or in the second column if his answer is "I don't know". If he finds an approximate answer ("about..."), then he can write this answer in the third column. If he finds the upper or lower limits (or both) to the answer, then he can use the fourth column to write down this answer. (This column is divided into two parts with a dotted line and a heading of "More than...Less than".) Finally, he can use the fifth column if he wishes to give an exact answer (which, as we mentioned earlier, does not exist for problem questions).

PART II

Part II of the test asks the subject to look at the questions to which he did not give an exact answer, and decide what method he would use to find a more exact answer. He is asked to make an X in one of the following eight columns for each question mentioned before. The eight columns are headed as follows:

I already gave an exact answer	To find a more exact answer I would...							
	ask my teacher	ask my father or my mother	look in my school books	look in another book	talk about it with a friend	think longer	experiment with it	find out some other way (EXPLAIN)

The first column: "I already gave an exact answer" serves only for checking whether the subject understood the instructions for the second part of the test.

23.2.3 Instructions to Subjects

Following are the directions to be used when presenting the Sherbrooke Problem Facing Questionnaire to a group of children.

PRELIMINARY REMARKS. Begin with whatever introduction is appropriate to the group and testing situation.

Distribute the test booklets. Say: Go ahead and look through the booklet to see what's inside.

To encourage the students to look through the booklets freely, pretend not to pay attention to them during this time.

Have the students write their names and other required information on the cover of the test booklet.

SAY: Put down your pencils and read the instructions silently while I read them to you aloud.

READ THE INSTRUCTIONS.

We'll do the first example together...

READ EXAMPLE 1 TWICE.

SAY, WHILE POINTING TO THE APPROPRIATE COLUMN: If your answer is "I haven't learned that", mark an X in the first column. If your answer is "I don't know", mark an X in the second column.

If you find an approximate answer, write this answer in the third column, under the word "about".

If you find that the answer is more than a certain number or less than another number, use the fourth column, the one divided into two parts:

If you find that the answer is more than a certain number, write that number in the column below the words "more than". If you find that the answer is less than a certain number, write that number in the column below the words "less than". You may use either one part of this column or both. And if you want to give an exact answer, write it in the last column.

Go ahead and write your answer and I'll look to see if you've understood. Verify. Say: "Good" to students who have put an X or an answer in a correct column...Help those who have marked an X where they should have written an answer etc. Ignore arithmetical mistakes.

Proceed to the second example.

READ EXAMPLE 2 ONCE.

Treat examples 2 and 3 in the same way as example 1.

After doing the three examples, say: Are there any questions? Does everyone understand how to write the answers? Now, turn the page and work through until you reach the end of the booklet. You may work in the space beneath each question or on the left hand page.

POINT OUT. You may use all the time you wish; when you finish, please raise your hand.

When each pupil finishes the test, give him part 2 and tell him: Fill in the top, read the beginning and do what it says to do.

Open each student's test booklet 1 at the first page of questions so that he will not treat the examples.

Explain the instructions to Part 2 to those students who do not understand.

Collect both booklets as each student finishes.

23.2.4 Scoring

The scoring of the Sherbrooke Problem Facing Questionnaire begins with a check whether the subject has understood the instructions of Part I (or has the minimal intelligence necessary for the test), that is, to determine if it is worth scoring. For this, the three questions which require exact answers have to be checked. The criterion is that the subject has at least two out of the three control questions answered in the fifth column.

Here as well as in the whole scoring process only the place of the answer is important. Whether the answer is really reasonable or not is not a point to be judged during the scoring process. This simplifies the scoring system and, more important, is in keeping with the overall purpose of the test: we wish to measure the subject's attitude towards the problems (whether he tries to answer the questions), not his knowledge or intelligence (how well he can answer the questions).

The score of Part II is mainly the difference between the number of the answers given in columns 5,6,7,8 and in columns 1,2,3. Because we did not use this part of the test in any of the studies described we do not intend to discuss it in detail here.

There are three possible ways at present for scoring Part I of the test.

(i) The Unweighted Scoring System

The simplest way is to count how many answers the subject gave in Column 1 and 2 (called Category 1: Avoidance) and separately how many answers the subject gave in Column 3 and 4 (Category 3: Problem-Facing). If f_1 is the number of answers in Category 1 and f_3 is the number of answers in Category 3, than

$$t = f_3 - f_1 \qquad \text{Eq. 23-1}$$

is the total problem facing score of the subject. (The answers in Column 5-Category 2: Neutral, do not count.)

(ii) Weighted Scoring System

The Weighted Scoring System* makes use of the weights derived from the item analysis described in 23.2.5. In this system, we began with the initial weights -1, 0, +1 for Categories 1, 2, and 3 respectively. Then, using an iteration process, we arrived at the final weights [ω(i,k), where i = 1, 2, 3 stands for the categories and k = 1, 2, ...22 stands for the questions]. The total score (T) is then simply the sum of the weighted answers:

---

* Designed by B. Connes.

$$T = \sum_{k=1}^{22} \sum_{i=1}^{3} \omega(i,k) \qquad \text{Eq. 23-2}$$

where ω(i,k) = the score in the appropriate column of Table I, if the subject gave his answer to question k in category i.

ω(i,k) = 0, if the subject did not give his answer to question k in category i.

(iii) Normalized Weighted Scoring System

Although the T score is sufficient for most purposes, for easier comparison with other tests and for better understanding of the test results (especially for teachers, who may be more familiar with IQ-type scales) the T-scores have been converted into an IQ-equivalent scale*. These new scores are then called normalized T scores (NT).

23.2.5 Statistical data

Any test in its fully developed form must fulfill two conditions:

1. It must really measure the trait it is intended to measure, and
2. It must give reliable data; that is, remeasuring the trait under similar conditions for the object or individual concerned we have to obtain the same (or very similar) results (Magnusson, 1966).

---

* IQ-equivalent scale is a scale like the Wechsler Intelligence Scale (WISC; Wechsler, 1949) which has a normal distribution, a mean of 100. and a standard deviation of 15. (For more details see e.g. Magnusson, 1966, pp. 235-238.)

The first condition (a test should be meaningful) is usually referred to as *validity*, while the second (a test should be reproducible) is termed *reliability*. It is quite difficult to satisfy these two conditions because there is a certain contradiction between them: For example, the validity of a test can be increased by making the test more heterogeneous, while for high reliability the test has to be homogeneous.

For some purposes reliability could be the more important objective for others validity. Generally a compromise between the two is the best solution.

Both the reliability and validity of a test depend on the properties of the individual items which make up the test.

The process of choosing those items from the lot which give the biggest contribution to the test's reliability and validity is called *item analysis*.

In the following pages we describe the collection these most important statistical data for the Sherbrooke Problem Facing Questionnaire (SPFQ).

RELIABILITY

*Theoretical considerations**

Reliability means that the result obtained from the instrument on a certain occasion under certain conditions should be reproducible.

There are several possible *sources of error variance*, e.g.

1. The *administration* of the test: Differences due to the behaviour of the test administrator, the interaction between the individuals testing and those being tested, physical environment, instruction.

---

* For a more detailed description see e.g. Stanley (1971) or Magnusson (1966).

2. The scoring of the test: If the scoring system is not entirely mechanical then there will always be some differences between scores given by two different scorers.

3. Fluctuation on the variable the test is supposed to measure; some characteristics vary within a certain individual in time very much.

4. Memory effect: If we repeat a test the result of the second administration will be influenced by the fact that the subject remembers - more or less, depending mainly on the time between the two testings - his responses in the first administration.

There are a number of different methods available for estimating reliability. The major procedures are the following:

1. Parallel tests

Administration of two parallel forms under specified conditions, and correlation of the resulting scores.

The reliability coefficient obtained is usually called equivalence coefficient.

2. Test-retest method

Repeated administration of the same test form or testing procedure, and correlation of the resulting scores.

The reliability coefficient obtained is usually called precision and stability coefficient.

3. Split-half method

Subdivision of a single test into two presumably parallel groups of items (half-forms a and b), each scored separately, and correlation of the resulting two scores. This yields an estimate of the reliability coefficient of the half-test, say r_{ab}. The reliability coefficient of the whole test is estimated by the Spearman-Brown prophecy formula (Eq. 23-3).

$$r_{a+b} = \frac{2\,r_{ab}}{1 + r_{ab}} \qquad \text{Eq. 23-3}$$

where r_{a+b} is the reliability coefficient for the whole test.

4. Analysis of the covariance among individual items, and determination of the true-score and error variance therefrom, which encompasses several seemingly diverse procedures.

As these methods take into consideration different combinations of the possible sources of error mentioned different methods are preferable for different purposes.

Reliability scores based on a French Canadian population

In the frame of the Sherbrooke study we used the following procedure to define the test's reliability: Form A and Form B were administered within 18 days (June 1, 1972 - June 19, 1972). The results of these tests will be referred to as A_1 and B_1; and the combination of the results as $(A_1 + B_1)$. About half a year later within a week (November 29, 1972 - December 6, 1972), Form A and Form B of the test were administered a second time to the same population. The results of these tests will be referred to as A_2 and B_2 and their combination as $(A_2 + B_2)$. The item analysis of Form $(A_1 + B_1)$ proved some items to be inadequate (see page 262-266). The remaining items were arranged in new forms S^4 and S^5. (S^4 contained those items which remained after the item analysis of grade 4 results, and S_5 those items which remained after the item analysis of grade 5 results.)

In what follows $S_1{}^4$ and $S_1{}^5$ stand for the results derived from $(A_1 + B_1)$ using those items involved in S^4 and S^5. $S_2{}^4$ and $S_2{}^5$ stand for those results derived from $(A_2 + B_2)$ using those items involved in S^4 and S^5.

In the calculation of the reliability coefficients we used both scoring systems available.

The results are shown in Table 23-1.

The Sherbrooke Problem Facing Questionnaire (French Canadian Form) in this study showed satisfactory reliability (especially with the split-half method).

VALIDITY

In general, the validity of a method is the accuracy with which meaningful and relevant measurements can be made with it, in the sense that it actually measures the traits it was intended to measure (Magnusson, 1966 p. 123). Briefly, the question of validity is whether the test really measures what it has been constructed to measure.

The key problem in testing is to establish the validity of the instruments and techniques developed. According to the American Educational Research Association (1955) and the American Psychological Association (1959) any testing instrument has to be evaluated using one or more of the four* types of validity techniques:

1. Predictive validity: A measurement of how the test predicts a certain outcome after a given time.
2. Concurrent validity: A measurement of how close the results of the test are to a certain criterion variable.
3. Content validity: An estimate of the extent to which the test covers the field it is intended to measure;
4. Construct validity: An evaluation of the extent to which test results support hypotheses based upon a certain theory or model.*

---

* Evaluation generally stresses content validity.

Table 23-1

Reliability coefficients of the Sherbrooke Problem Facing Questionnaire (SPFQ) French Canadian Form

FORM / METHOD		A				B				A + B	S^4				S^5			
	Based on Testing Period	Initial weights	Final Weights ω_1	Final Weights ω_2	ω_1/ω_2	Initial Weights	Final Weights ω_1	Final Weights ω_2	ω_1/ω_2	Initial Weights	Initial Weights	Final Weights ω_1	Final Weights ω_2	ω_1/ω_2	Initial Weights	Final Weights ω_1	Final Weights ω_2	ω_1/ω_2
Parallel Tests Equivalence Coefficient	First	A_1B_1 .725	A_1B_1 .642	—	—	A_1B_1 .725	A_1B_1 .642	—	—	—	—	—	—	—	—	—	—	—
	Second	A_2B_2 .660	—	A_2B_2 .686	—	A_2B_2 .660	—	A_2B_2 .686	—	—	—	—	—	—	—	—	—	—
Test-Retest Method Stability Coefficient	-------	A_1A_2 .622	A_1A_2 .564	—	A_1A_2 .523	B_1B_2 .666	B_1B_2 .538	—	B_1B_2 .591	(A_1+B_1) (A_2+B_2) .660	$S_1^4\ S_2^4$ .726	$S_1^4\ S_2^4$ .515	—	$S_1^4\ S_2^4$ .341	$S_1^5\ S_2^5$ .682	$S_1^5\ S_2^5$ .658	—	$S_1^5\ S_2^5$ .560
Split-Half* Method	First	$A_1^o\ A_1^e$ .789	$A_1^o\ A_1^e$ .792	—	—	$B_1^o\ B_1^e$ .813	$B_1^o\ B_1^e$ .587	—	—	$(A_1+B_1)^o$ $(A_1+B_1)^e$ .868	$S_1^{4o}\ S_1^{4e}$ .882	$S_1^{4o}\ S_1^{4e}$ .878	—	—	$S_2^{5o}\ S_2^{5e}$ — .844	$S_1^{5o}\ S_1^{5e}$.[illegible]57	—	—
	Second	$A_2^o\ A_2^e$ .846	—	$A_2^o\ A_2^e$ .823	—	$B_2^o\ B_2^e$ .7[illegible]1	—	$B_2^o\ B_2^e$ .794	—	$(A_2+B_2)^o$ $(A_2+B_2)^e$ .907	$S_2^{4o}\ S_2^{4e}$ .874	—	$S_2^{4o}\ S_2^{4e}$ .836	—	$S_2^{5o}\ S_2^{5e}$ .896	—	$S_2^{5o}\ S_2^{5e}$.[illegible]	—

All correlation coefficients are significant at .1% level.
Form A, B and A + B results are based on 104 grade 4 and grade 5 subjects.
Form S_4 is based on 53 grade 4 and Form S_5 on 51 grade 5 subjects.

* The split-half coefficients are corrected with the Spearman-Brown prophecy formula:

$r_{a+b} = \frac{2r_{ab}}{1+r_{ab}}$ where r_{ab} = the correlation between the two halves of the test and r_{a+b} = the estimated reliability of the whole test

o,e refers to the split between even and odd numbers.

It should be noted that useful as a combined assessment according to all four criteria would be only very few tests have in fact been validated by all four methods.

Predictive and Concurrent Validity

Predictive validity and concurrent validity are expressed as a validity coefficient: The correlation between the test results and some predicted outcome (e.g. ratings after a certain period of employment) in the former case, and correlation with some criterion variable available at the same time (e.g. clinical diagnosis of brain damage) in the latter case.

At present the Sherbrooke Problem Facing Questionnaire (SPFQ) is too new to give data about its predictive validity. As for its concurrent validity, unfortunately no single method provides a criterion variable for attitude towards problems.

Content Validity

Content (and also construct) validity cannot be expressed as a simple correlation coefficient. To determine content validity, competent judges should define the appropriateness of the particular test to the particular problem or situation (Bloom, 1970). Of course the definition of a "competent judge" remains an open question and depends on the human characteristics being measured. For evaluation of the psychological effect of different teaching methods - a purpose for which the Secret Writing test is recommended and which stresses the content validity of the tests being used - the judge is likely to be a person with considerable competence in the educational system to be evaluated and with considerable training and experience in the dynamic theories used in differential psychology.

In the particular case of the Sherbrooke Problem Facing Questionnaire (SPFQ) we could gain useful information concerning its content validity by studying how well the questions cover the field of cognitive problems.

We hope to carry out a study to help clarify this question. But although the results of this kind of study can contribute to the validation of the test, they can not guarantee its content validity. They do not prove that the test is appropriate to the field it is intended to measure. They certainly cannot reveal whether the things being measured are those that ought to be measured. Nobody but the educator or psychologist who wishes to test the children can decide whether a particular test is adequate for his particular purpose. To answer this question the examiner must "take" the test himself, putting himself in the place of the pupil and thinking about the test the way he believes the pupil will think about it. Through introspection he should attempt to determine what mental processes are needed to complete the test and whether these mental processes are important for the testing purpose (e.g. whether these mental processes are the ones the evaluated schools want to develop through instruction).

Although this kind of evaluation has the many disadvantages of any other subjective method and is a difficult, painstaking and time-consuming process, it is the only way in which the tester can discern clearly the most important of the differences prevailing among available tests. Considering the importance of the inferences the tester will later draw from the test results, and the influence that the test may later exert upon the individuals tested or upon the instruction and guidance in the school in which the test is used, the task is certainly worth doing to the best of the tester's ability.

Construct Validity

Construct validity is a fairly new concept launched by a technical committee appointed by the American Psychological Association in 1954. Although some psychologists have still rather critical views on this concept (see, e.g., Bechtoldt, 1959) it is fairly widely used, especially for tests measuring traits for which external criteria are not available. (For example, educational assessment - the attempt to assess the characteristics of individuals in relation to a particular environment or task - has characteristically emphasized construct validity, since it goes to more elaborate lengths to use theories or models to guide it in particular assessment situations. It is these theories or models that make it possible to use construct validity, and it is the very complexity of the data collection and analysis required in assessment that make it necessary to have a theory or model to guide the workers through the intricacies of an assessment process. See Bloom, 1970.)

The procedure for testing construct validity is an application of the deductive method - currently applied in all scientific research - to the evaluation of the accuracy of predictions based upon a test. This evaluation follows the classical procedure: theory - deduction - hypothesis - experimental testing - data which falsify or verify the hypothesis (Magnusson, 1966).

Construct validity can be tested in several different ways, some of which are the followings (Magnusson, 1966; for more ways see, e.g., Cronbach and Meehl, 1955):

1. The study of differences between groups which should differ according to the theory for the variable.

2. The study of how the test results are influenced by changes in the individuals or the environment which, according to the theory, should or should not influence the individuals' position on the continuum.

3. The correlation between different tests which are assumed to measure the same or very similar variables.

ITEM ANALYSIS

The need for item analysis arises from a very practical problem in testing: Because he usually has only a very limited time available for testing and scoring, the psychologist or educator is looking for a measurement as reliable and as valid as possible but with as small a number of items as possible. To fulfill this need those who construct the test have to use a method which provided data about the likelihood of each item's increasing the reliability and the validity of the test.

i) An item contributes to a test's reliability by measuring the same sort of true scores as the other items in the test. This is usually measured by the correlation between the item and test. The reliability index is a product of this and the standard deviation of the item.

ii) An item contributes to the test's validity if it measures the same sort of true score as the criterion measurement. This is usually measured by the correlation between the item and the criterion. The validity index is a product of this and the standard deviation of the item.

The objective of the item analysis is then to choose those items with the highest reliability and validity index at the same time.

The methods adopted by us

As stated earlier, at present no single method yields a criterion variable for validating a problem-facing test. In our approach we therefore confined ourselves to the assessment of the reliability of the items.

The essence of the item analysis adopted by us (and based on the design of the Sherbrooke Item Bank, Connes, 1973*) is the following:

The method proceeds by successive approximations. In what follows the subscript i runs over the set of questions and the subscript j over the set of subjects.

$k(i,j)$ is the response of the individual j to the question i.

$\left[k(i,j) = 1, 2 \text{ or } 3\right.$ depending on which category this answer falls; see $\left.5.1\right]$

$\omega(i,k\left[i,j\right])$ is the weight or value of response k for question i

(initially -1, 0, or +1 depending on k; see 23.2.4)

1) For each subject j, a total score $\left[T(j)\right]$ can be computed:

$$T(j) = \sum_{i=1}^{I} \omega\left[i, k(i,j)\right] \qquad \text{Eq. 23-4}$$

where I = number of questions.

2) The mean (m) and standard deviation (σ) of T(j) are

$$m = \frac{1}{J} \sum_{j=1}^{J} T(j) \qquad \text{Eq. 23-5}$$

---

* This item analysis method was designed by B. Connes.

$$\sigma = \sqrt{\frac{\sum_{n=1}^{J} T(j)^2}{J} - m^2} \qquad \text{Eq. 23-6}$$

where J = number of subjects.

3) Standardize the distribution of T (j), giving

$$T'(j) = \frac{T(j) - m}{\sigma} \qquad \text{Eq. 23-7}$$

for j = 1, 2, ... J.

4) For each question, read just the initial weights $\omega(i,k)$ using the formula:

$$\omega'(i,k) = \frac{\sum_{i=1}^{J} T(j)\, C(i, j, k)}{P(i,k)\, J} \qquad \text{Eq. 23-8}$$

$$\text{where } C(i,j,k) = \begin{cases} 1, & \text{if subject } j \text{ gave answer } k \text{ to question } i. \\ 0, & \text{if not.} \end{cases}$$

$$\text{and} \quad P(i,k) = \frac{1}{J} \sum_{j=1}^{J} C(i, j, k) \qquad \text{Eq. 23-9}$$

Thus P(i,k) is simply the proportion of subjects giving answer k to question i.

If the new weights $\omega'(i,k)$ for question i are not in the expected order

$$\omega'(i,1) \leq \omega'(i,2) \leq \omega'(i,3)$$

then the question must be at least temporarily eliminated by giving it the weight 0 for all answers.

5) If the difference between the new $\omega'(i,k)$ and the old $\omega(i,k)$ is less than a given level of convergence for all i and k, then the item analysis and weighting processes are finished; if not, begin again at step 1.

Results of the item analysis

The item analysis based on more than 100 subjects gave us quite satisfactory results;

i) it served us with the appropriate weights of the different answers to the different questions, practically between the same limits as those of the initial weights: -1 and +1.

ii) The intercorrelation between the items was between low (0) and moderate (0.4), calculated with Pearson coefficients.

iii) The correlations between items and test were moderate (mostly between 0.3 and 0.6, all significant at 0.01% level), calculated with triserial correlation (Japsen, 1946; Jenkins, 1956).

iv) It showed that the items have a large enough range of difficulty level.

v) It showed that more than 24 out of the 36 items proved useful.

The Iowa Study

The American Version of the Sherbrooke Problem Facing Questionnaire Form (A + B) - a direct translation of the French Canadian Version - was administered in December 1972 in three elementary schools in Cedar Falls, (Iowa, U.S.A.). About 190 fifth graders participated in this study. Some of the results of the Iowa study were the following:

i) The item analysis served us with the appropriate weights for the different answers to the different questions (practically between the same limits as those of the initial weights and as in the French Canadian Version between -1 and +1. These weights however were not used in the New York Study, the purpose of which was to evaluate students with two different cultural and language backgrounds.

ii) The intercorrelations between the items were similar to those Canadian Version and ranged from low (0) to moderate (0.5).

iii) The triserial correlations between the items and the total score derived from the weighted scoring system (except in one case) were between moderate (0.3) and fairly high (0.9).

iv) The difficulty levels of the items had a fairly good distribution in this grade.

v) As in the case of the French Canadian Version more than 24 out of the 36 items proved to be useful.

vi) The distribution of the T scores was closer to the standard normal distribution (the T scores were also converted into an IQ type of scale).

vii) Table 23-2 shows that (except for some very low correlations) we did not find any connection between the SPFQ and achievement in the Stanford Achievement Test or between SPFQ and the Kuhlman-Anderson intelligence test. These results support the hypothesis that attitude towards problems is independent of school achievement and also of intelligence.

23.3 The Sherbrooke Study

23.3.1 Method

Since the method is fully described in the preceding sections (see 23.2.2, 23.2.3 and 23.2.4) we need only add that for the comparison of the experimental and control classes the data of the first testing in June 1972 were used. The data were obtained by administering the French Canadian Forms A and B in different sessions with 4 to 6 days in between.

Table 23-2

Correlation between the Sherbrooke Problem Facing Questionnaire (SPFQ), Form R American Version (using weighted scoring system) and the Stanford Achievement Test (Kelley et al, 1964) and Kuhlman-Anderson intelligence test with Pearson correlation coefficients, based on grade 5 elementary school children from Cedar Falls, Iowa, U.S.A.

REFERENCE TEST / SPFQ	STANFORD ACHIEVEMENT TEST										KUHLMAN-ANDERSON INTELLIGENCE TEST (IQ)
	WORD MEANING	PARAGRAPH MEANING	SPELLING	LANGUAGE	ARITHMETIC COMPUTATION	ARITHMETIC CONCEPTS	ARITHMETIC APPLICATION	SOCIAL STUDIES	SCIENCE	MEDIAN	
TOTAL SCORE (T)	0.024	0.088	0.039	0.079	0.166*	0.074	0.106	0.146*	-0.019	0.118	0.136
NORMALIZED TOTAL SCORE (NT)	0.030	0.095	0.046	0.082	0.167*	0.069	0.115	0.154*	-0.015	0.126	0.148*
NUMBER OF SUBJECTS	158	158	160	160	159	159	159	160	160	159	163

*Significant at 5 percent level.

23.3.2 Treatment of data

Both unweighted and weighted scores had been used to compare experimental and control groups (see 23.2.4). The item analysis was carried out in grades 4 and 5 separately. Experimental-control differences were examined by t tests.

23.3.3 Expectation

Students involved in the Dienes method of learning mathematics should show a more favourable attitude towards problems - facing rather than avoiding them.

23.3.4 Results

Whether the raw scores (Table 23-3)* or the T values (Table 23-4) are taken as a basis of comparison the results are the same. (Fig. 23-2 shows the cummulative frequency distribution of the T scores.)

The students of the experimental class in grade 5 had a significantly better attitude towards problems. In grade 4 no such a difference was revealed by our data. The findings support our hypothesis at least in grade 5, where children have been exposed to the Dienes method of mathematics learning for a longer time than in grade 4.

23.4 The New York Study

Here we attempted to answer the question:

"What is the position of Fleming school children in their attitude towards problems compared with the children in Sherbrooke and another group of American students?"

---

*Tables 23-3 to 23-5 will be found at the end of this chapter.

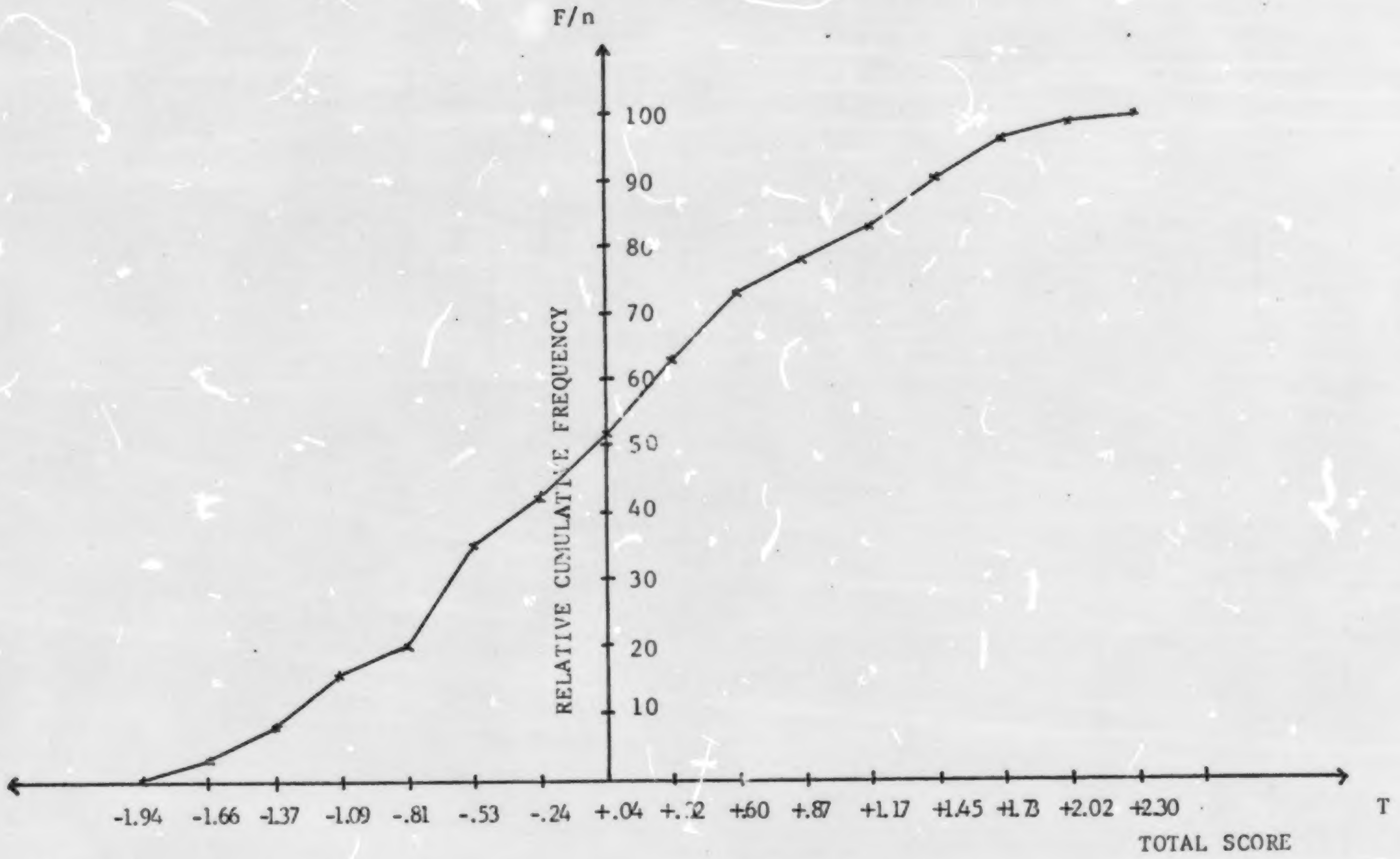


Fig. 23-2. The cumulative frequency distribution of the total score (T) in the Sherbrooke Problem Facing Questionnaire (SPFQ French Canadian version), based on a sample of 104 grade 4 and 5 students. (mean = 0; S.D. = 1; numbers referring to scale units are rounded to the nearest two decimal places.)

23.4.1 Method

The method is described in detail in sections 23.2.2 to 23.2.4. In this study Form (A + B) was used, but the questions were arranged for every student in a random order. We gave all the 44 questions in one sitting because former studies have shown grade 4 and 5 children can handle that many problems without showing symptoms of fatigue.

23.4.2 Treatment of data

For scoring the unweighted system was used. The differences between Sherbrooke, New York and Iowa data were determined by t tests.

23.4.3 Results

In our comparison of the results obtained in the New York School we used data obtained at Cedar Falls, Iowa, and at Sherbrooke, Quebec. The Sherbrooke data consisted of the combined scores obtained from an experimental and a control school, and the Iowa data were taken from a school that had not been exposed to modern mathematics of the Dienes type at all.

An overview of the results is given in Table 23-5:

i) Looking at the means in grade 5, the position of the New York students is below that of the Sherbrooke students (t = 2.55) and also below that of the Iowa students (t = 5.33).

ii) Looking at the means in grade 4, however, there is no evident difference between the New York and the Sherbrooke students (t = 1.09, not significant). No grade 4 students were tested at Iowa.

iii) Comparing the 4 classes of the New York School with each other we notice that the yellow classes are superior to the red classes. In grade 5 we have means of -1.60 and -7.27 respectively, while in grade 4 they are 9.38 and -4.07 respectively.

The mean of 9.38 for the yellow grade 4 is the highest we recorded anywhere, but without further investigations its interpretation must remain a matter for conjecture.

Table 23-3
The mean differences between experimental and control classes in the Sherbrooke Problem-Facing-Questionnaire (SPFQ), based on raw scores (weights 1,2,3)

CLASS	Statistics Group	N	$\bar{X}$	SD	t
4	exp.	19	73.79	11.70	-0.67
	contl.	17	76.65	13.45	
5	exp.	22	84.32	14.03	2.83***
	contl.	21	73.76	10.14	
4+5	exp.	41	79.44	13.90	1.522
	contl.	38	75.05	11.66	

*** significant at the .01% level.

Table 23-4

The mean differences between the experimental and control classes in the Sherbrooke Problem Facing Questionnaire (SPFQ). The comparison in grade 4 is based on T values of Form S_4 and in grade 5 on T values of Form S_5.

Grade	Statistics Groups	N	$\bar{X}$	SD	t
4	exp.	19	-0.043	1.028	-0.903
	contl.	17	0.255	0.948	
5	exp.	22	0.245	1.094	2.487**
	contl.	21	-0.491	0.822	

** significant at the .05 percent level.

Table 23-5

The results of the Sherbrooke, Iowa and New York Studies with the Sherbrooke Problem Facing Questionnaire (SPFQ), unweighted scoring system.

STATISTICS / STUDIES			MEAN X	NUMBER of SUBS	STANDARD DEVIATION SD	F	t (two sided)
NEW YORK	"ROUGE"	4	-4.07	15	12.51		
		5	-7.27	15	7.49		
	"JAUNE"	4	9.38	16	14.90		
		5	-1.60	15	14.91		
	TOTAL	4	2.87	31	15.19		
		5	-4.43	30	11.95	1.53	5.33**
IOWA		5	8.60	182	14.27	1.48 1.24	1.09 2.55*
SHERBROOKE***		4	-0.86	36	12.47		
		5	3.16	43	13.28	1.11	2.24*

* significant at 5 percent level

** significant at 0.1 percent level

*** the experimental and control school together

Chapter 24

Summary

1. Budapest study

Summarizing all the results obtained in the Budapest study, it could be said that evidence was found more or less supporting all hypotheses except hypothesis 4.

There are several possible explanations for the lack of results in hypothesis 4 (greater concentration of the children's attention). Besides the simple explanation that the mathematics lesson may be as interesting for the students in the control classes as in the experimental classes, there exists another equally valid possibility. The cause of the Toulouse-Pieron test result differences after a school lesson is probably an increment in the students'arousal. But this can be a result either of an internal interest (unintentional attention on work or play, etc) or an external pressure (intentional attention on work, avoidance of recitation, etc.). Then the real failure of this method is in differentiating between these two possibilities.

Looking now at the other hypotheses, the students in the experimental classes showed at least equal knowledge of "traditional" mathematics to that of those in the control classes, while the experimentals had in addition a fairly considerable knowledge of "modern" maths. This difference in the amount of information the students absorb in one mathematics lesson does not cause greater fatigue. In fact there was some evidence that the experimental method is less tiring. (This statement was supported not only by the objective measurement of fatigue, but also by the semantic differential method in which the mathematics lesson was rated as less tiring -- or more refreshing — by the students in the experimental classes.)

We found significantly better results (at least in grade 5) in the experimental class on the learning ability test as well as a higher level of creativity as measured by several of its parameters. The gain in creativity was accompanied by a loss of rigidity in all the grades examined. A greater agreement was found in the students' judgement, and a more favourable attitude especially towards the mathematics lesson was measured in the experimental classes. The examination of the students' anxiety showed that (at least in grade 4) the students in the experimental class were less anxious, and the measurement of n Ach showed that they had a greater need to achieve. By combining the results of the latter two tests it may be concluded that the students in the experimental classes have a higher achievement motivation.

In the whole process of evaluation some serious mistakes were made: technical errors (as in the measurement of finger tremor), organizational problems (as in the measurement of modern mathematics knowledge) and even methodological mistakes (as in the measurement of the concentration of attention mentioned above). It is particularly unfortunate that we did not examine well enough the interrelations between the measurements we used, but the few results we got in this field show a slight connection: the IA scores had a 0.35 correlation with the I.Q. scores in grade 5, the creativity scores had a low negative correlation with the perseveration index (with a maximum in grade 5 originality scores:-0.32, see Table 20-1).

2. Sherbrooke Study

In spite of the greater effort expended on this Study the differences that emerged were less clear-cut than those obtained in Budapest. Two possible reasons might account for this discrepancy:

i) Choosing the experimental and control groups from different schools may have increased the "statistical noise", i.e. the disturbing factors.

ii) Whereas in the Budapest Study the control classes were genuinely "virgin" to the Dienes-type methods of learning these methods had penetrated into almost all the Sherbrooke schools by the time the evaluation was carried out, with the result that our control classes were probably affected to some degree.

Looking at details it appears that the two tests which failed to produce any significant differences were Demangeon's Anxiety Test and Zazzo's Perseveration Test. The Learning Ability Test (Secret Writing) also failed to reveal a significant difference between the experimental and control classes, though in this case grade 5 children showed a significant gain against the children of grade 4 in the experimental but not in the control classes. Again, Osgood's Semantic Differential Method for measuring attitude towards the mathematics lesson and Cattell's Children's Personality Questionnaire gave mostly controversial results.

The children in the experimental classes proved to be less competent at traditional mathematics (Iowa Tests of Basic Skills), especially in the part of the Test which stresses computational aspects.

On the other hand the children in the experimental classes gained a great deal of knowledge in modern mathematics; their creativity became distinctly higher than that of the control groups (Torrance Test of Creative Thinking), and they developed a positive attitude towards facing new cognitive problems rather than avoiding them (Sherbrooke Problem Facing Questionnaire).

These gains are regarded by us as highly important, and it is our intention to look for similar gains in other directions in the future.

3. New York Study

The evaluation study of the Fleming School in New York served two different purposes. The evaluators' interest was the extension of the evaluation program of the Dienes Method of Mathematics Learning which covered at that point two projects, one in Budapest (Hungary) and one in Sherbrooke (Québec), both now having completed their main phase. Since in the Sherbrooke study development of new tests was a principal target of research aimed at filling a gap in methods available for evaluation in the field of progressive education, data on these tests from another cultural background were welcome.

The main purpose however, was determined by the director of the Fleming School herself. She wanted answers to 2 questions:

(i) What could be said about the children of her school through objective measurement, on aspects other than intelligence and achievement?

(ii) What was their standard in mathematics in relation to the school in Sherbrooke (Québec) which was closely related to the creator of the Mathematics Program she was using?

As the results of the preceding chapters show, the children of grade 4 and 5:

(i) Have a generally positive attitude towards their teacher, their school and their mathematics lesson.

(ii) Show a high standard of learning ability compared with students in the Sherbrooke school as measured by a recently developed test.

(iii) Are quite different in their attitude towards problems in terms of their class means. (The evaluators found no plausible explanation for this outcome, since the test situation was the same in all classes and the test itself has a good reliability.)

(iv) Obtained somewhat better scores in achievement tests in Modern Mathematics.

(v) Are doubtless superior in their achievement in Traditional Mathematics to the students in Sherbrooke.

It is advisable to interpret the differences between the two schools under investigation only through differences in socio-economical level, intelligence level and cultural-linguistic environment. Quite apart from this, the results also show that the children of the Fleming school seem to be happy with their work and are doing well at both modern and traditional mathematics.

4. Conclusion

In this book we have tried to describe the nature and objectives of educational evaluation, and to bring out the large number of difficulties by which those who embark on it are faced. In particular we have recorded our attempts to evaluate the psychological effects of the Dienes method as a way of teaching modern mathematics at the elementary levels. In the course of more than four years and at the cost of innumerable working hours we carried out three evaluation studies in two different continents in three different countries. A considerable amount of manpower was used for planning, developing and adapting tests, also for testing the children, scoring answer sheets and preparing statistical analyses. To this must be added the large amount of money required for the purchase of materials and the operation of machinery. Again, while running the tests we were forced to disturb the teachers and children in their work on many occasions.

This makes it imperative that we ask ourselves in all sincerity whether the effort is justified by its results. Unfortunately no ready yes-or-no answer is forthcoming. Our three studies had a particular as well as a more general objective. On the particular side we wanted to produce answers to the many questions which the children, parents, teachers, educators, psychologists and administrative leaders involved have posed concerning the very interesting and important pedagogical experiment called the Dienes method. On the general side we sought to develop a set of instruments sensitive enough to measure the psychological effects of any new kind of learning situation.

As far as our particular objective is concerned we are able to provide some answers: The children in the Hungarian experimental classes were not more tired than those of the control classes; the children in the New York experimental school, while still possessing good knowledge of the traditional topics, gained much knowledge in the ever-growing and important field of modern mathematics etc. In other directions, such as the children's powers of generalization or their creativity, our conclusions are far less definite. The usefulness of these answers must be left to be judged by those who asked the questions. So far the most favourable verdict has been given in Budapest, where our investigations have helped teachers in their everyday work and the Ministry of Education has given the OPI Mathematics Project its full support, to develop into a nationwide scheme.

Our general objective has not been attained and may, in the opinion of some co-operating psychologists, be unattainable because of the magnitude of the tasks involved. We were unable to construct the instruments as projected, but if we can convince the reader that the task of searching for their construction remains an urgent one then our effort will have been worth while.

It is our intention to continue with our work in educational evaluation, and we believe that the construction of new tests especially designed for the evaluation of a particular project is a useful tool in this work. We are equally aware that the task we have set ourselves is challenging as it is ardous. If the reader shares our awareness and feels that he can offer suggestions for help let him rest assured that these will always be gratefully received.

BIBLIOGRAPHY

Alexakos, C.E. Predictive efficiency of two multivariate statistical techniques in comparison with clinical predictions. Journal of Educational Psychology, 5, 1966.

Allison, R.B. Learning parameters and human abilities. Office of Naval Research Technical Report. Princeton, N.J.: Educational Testing Service, 1960.

Alpert, R. and Haber, R.N. Anxiety in academic achievement situations. Journal of Abnormal Social Psychology, 61, 1960.

American Educational Research Assoc. Technical Recommendations for Achievement Tests. Washington, D.C.: AERA, 1955.

American Psychological Assoc. Technical recommendations for psychological tests and diagnostic techniques. Washington, D.C.: HPA 1959.

Amidon, E. and Flanders, N.A. The effect of direct and indirect teacher influence on dependent-prone students learning geometry. Journal of Educational Psychology, 52, 1961.

Amthauer, R. Intelligenzstrukturtest. Göttingen, 1953.

Anderson, R.C. Individual differences and problem solving. In R.M. Gagné (ed.): Learning and Individual Differences: A Symposium of the Learning Research and Development Center, University of Pittsburg. Columbus, Ohio: Charles E. Merrill Publishing Co., 1967.

Anderson, R.C., Faust, G.W., Roderick, M.C., Cunningham, D.J., and André, T. Current Research on Instruction. Englewood Cliffs, N.H.: Prentice-Hall, Inc., 1969.

Ashlock, R.B. Objectives in mathematics education: Can their attainment be observed? In R.B. Ashlock and W.L. Herman (eds.): Current Research in Elementary School Mathematics. London: The Macmillan Company, 1970.

Ashlock, R.B. and Herman, W.L. (eds.): Current Research in Elementary School Mathematics. London: The Macmillan Company, 1970.

Atkinson, J.W.(ed) Motives in Fantasy, Action and Society: A method of assessment and study. Princeton: D. Van Nostrand. 1958

Atkinson, J.W. An Introduction to Motivation. New York: D. Van Nostrand, 1964.

Atkinson, J.W. and Feather, N.T. (eds.): A Theory of Achievement Motivation. New York: John Wiley and Sons, 1966.

Atkinson, J.W. and Litwin, G.H. Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure. Journal of Abnormal Social Psychology, 60, 1960.

Atkinson, J.W. and O'Connor, P. Effects of ability grouping in schools related to individual differences in achievement-related motivation. In J.W. Atkinson and N.T. Feather (eds.): A Theory of Achievement Motivation. New York: John Wiley and Sons, 1966.

Atkinson, J.W. and Reitman, W. Performance as a function of motive strength and expectancy of goal attainment. Journal of Abnormal Social Psychology, 53, 1956.

Ausubel, D.P. Learning Theory and Classroom Practice: Bulletin No. 1. Toronto: The Ontario Institute for Studies in Education, 1967.

Ausubel, D.P. Fostering creativity in the school. In D.W. Brison (ed.): Accelerated Learning and Fostering Creativity. Toronto: The Ontario Institute for Studies in Education, 1968a.

Ausubel, D.P. Educational Psychology: A Cognitive View. New York: Holt, Rinehart, and Winston, Inc., 1968b.

Avital, S.M. and Shettleworth, S.J. Objectives for Mathematics Learning: Some Ideas for the Teacher: Bulletin No. 3. Toronto: The Ontario Institute for Studies in Education, 1968.

Bahrick, H.P., Fitts, P.M., and Briggs, G.E. Learning curves: facts or artifacts? *Psychological Bulletin*, *54*, 1957.

Barron, F. *Creativity and Psychological Health*. Princeton, N.J.: Van Nostrand, 1963.

Barkóczi, I. and Klein, S. Gondolatok az alkotóképességről és vizsgálatának egyes problémáirol. (Reflections on creativity and some problems in its measurement). *Magyar Pszichológiai Szemle*, *25*, 1968.

Beaudet, A. Recherche sur la créativité. *Thèse de $3^{\text{ième}}$ Cycle.* Université de Caen: Non publié, 1966.

Bechtoldt, P. Construct validity: a critique. *American Psychologist*, *14*, 1959.

Begle, E.G. and Wilson, J.W. Evaluation of mathematics programs. In E.G. Begle: *Mathematics Education: The 69th Yearbook of the National Society for the Study of Education*. Chicago: The University of Chicago Press, 1970.

Berlyne, D.E. *Conflict, Arousal, and Curiosity*. New York: McGraw-Hill, 1960.

Berlyne, D.E. Can and should learning be accelerated? In D.W. Brison: *Accelerated Learning and Fostering Creativity*. Toronto: The Ontario Institute for Studies in Education, 1968.

Berlyne, D.E. What next? (Concluding summary). In H.I. Day, D.E. Berlyne, and D.E. Hunt (eds.): *Intrinsic Motivation: A New Direction in Education*. Toronto: Holt, Rinehart, and Winston of Canada, 1971.

Berlyne, D.E. and Lewis, J.L. Effects of heightened arousal on human exploratory behaviour. *Canadian Journal of Psychology*, *17*, 1963.

Biczok, F. A felvételi vizsga specialis kérdései a tanari szakokon. (Some special problems of the entrance examination to the teachers colleges). Felsőoktatasi Szemle, 6, 1967.

Bieri, J. Complexity-simplicity as a personality variable and preferential behavior. In D.W. Fiske and S.R. Maddi (eds.): Functions of Varied Experience. Homewood, Illinois: Dorsey Press, 1961.

Biggs, J.B. Anxiety, Motivation, and Primary School Mathematics: Occasional Publication No. 7. London: Information Service of the National Foundation for Educational Research in England and Wales, 1962.

Biggs, J.B. Mathematics and the Conditions of Learning: A Study of Arithmetic in the Primary School. London: Information Service of the National Foundation for Educational Research in England and Wales, 1967.

Binet, A. and Simon, H. Application des methodes nouvelles au dianostic du niveau intellectuel chez les enfants normaux et anormauz d'hospice et d'école primaire. L'Annuaire Psychologique, 11, 1905.

Birch, H.G. The role of motivational factors in insightful problem-solving. Journal of Comparative Psychology, 38, 1945.

Bloom, B.S. (d.): Taxanomy of Educational Objectives: The Classification of Educational Goals: Handbook I: Cognitive Domain. New York: David McKay Co., 1956.

Bloom, B.S. Learning for mastery: evaluation comment I, no. 2. In Bloom, Hastings, Maddaus, et al.: Formative and Summative Evaluation of Student Learning. New York: McGraw-Hill, 1968.

Bloom, B.S. Some theoretical issues relating to educational evaluation. In R.W. Tyler (ed.): Educational Evaluation: New Roles, New Means: The 68th Yearbook of the National Society for the Study of Education. Part II. Chicago: The University of Chicago Press, 1969.

Bloom, B.S. Toward a theory of testing which includes measurement-evaluation assesment. In M.C. Wittrock and D.E. Wiley (eds.): The Evaluation of Instruction: Issues and Problems. New York: Holt, Rinehart, and Winston, Inc., 1970.

Bock, R.D. Contributions of multivariate experimental designs to educational research. In R.B. Cattell (ed.): Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally and Co., 1966.

Bodwin, R.F. The relationship between immature self-concept and certain educational disabilities. Doctors dissertation. Lansing: M.S.U., 1957.

Brickell, H.M. Appraising the effects of innovations in local schools. In R.W. Tyler (ed.): Educational Evaluation: New Roles, New Means: The 68th Yearbook of the National Society for the Study of Education. Part II. Chicago: The University of Chicago Press, 1969.

Brown, G.I. An experiment in the teaching of creativity. School Review, 1964, 72, 437-450.

Brownell, W.A. Arithmetical Abstractions: The Movement toward Conceptual Maturity under Differing Systems of Instruction. University of California Publications in Education: Vol. 17. Berkeley: The University of California Press, 1967.

Brownell, W.A. Conceptual maturity in arithmetic under differing systems of instruction. The Elementary School Journal, 69, 1968.

Brumlik, J. On the nature of normal tremor. Neurology, 12, 1962.

Bruner, J., Goodnow, J.J., and Austin, G.A. A Study of Thinking. New York: John Wiley and Sons, Inc., 1956.

Buros, O.K. (ed.): The Sixth Mental Measurement Yearbook. Highland Park, New Jersey: Gryphon Press, 1965.

Buros, O.K. (ed.): Personality, Tests and Reviews. Highland Park, New Jersey: Gryphon Press, 1970.

Burt, C. The appropriate uses of factor analysis and analysis of variance. In R.B. Cattell (ed.): Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally and Co., 1966.

Cahen, L. Comments on Professor Messick's paper. In M.C. Wittrock and D.E. Wiley (eds.): The Evaluation of Instruction: Issues and Problems. New York: Holt, Rinehart, and Winston, Inc., 1970.

Cambridge Conference on School Mathematics, Goals for School Mathematics. Boston, Mass.: Educational Services, Houghton Mifflin Co., 1963.

Carlson, H.B., Fischer, R.P., and Young, P.T. Improvement in elementary psychology as related to intelligence. Psychological Bulletin, 42, 1945.

Carmen, R.A. Programozott vektoralgebra (Programmed Vector Algebra). Budapest: Tankönyvkiadó, 1968.

Caron, A.Y. Curiosity, achievement, and avoidant motivation as determinants of epistemic behaviour. Journal of Abnormal Social Psychology, 67, 1963.

Castaneda, A. Effects of stress on complex learning and performance. Journal of Experimental Psychology, 52, 1956.

Castaneda, A. Supplementary Report: Differential position habits and anxiety in children as determinants of performance in learning. Journal of Experimental Psychology, 61, 1961.

Castaneda, A., McCandless, B.R., and Palermo, D.S. The children's form of the manifest anxiety scale. Child Development, 27, 1956a.

Castenada, A., McCandless, B.R., and Palermo, D.S. Complex learning and performance as a function of anxiety in children and task difficulty. Child Development, 27, 1956b.

Cattell, R.B. On the measurement of perseveration. British Journal of Educational Psychology, 5, 1935.

Cattell, R.B. Description and Measurement of Personality. London: G.G. Harrap, 1946.

Cattell, R.B. Children's personality questionnaire, test booklets, Champaign, Illinois: Institute for personality and Ability testing, 1963.

Cattell, R.B. (ed.): Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally and Co., 1966.

Cattell, R.B. and Cattell, A.K.S. Culture Fair Intelligence Test. Champaign, Illinois: Institute for Personality and Ability Testing, 1949.

Cattell, R.B. and Cattell, A.K.S. Epreuve d'intelligence sans apport culturel. Montréal: Institut de Recherches Psychologiques, 1967.

Cattell, R.B. and Scheier, I. The Meaning and Measurement of Neuroticism and Anxiety. New York: Ronald Press, 1961.

Cattell, R.B. Children's personality questionnaire, manual and Handbook, 1968. Champaign, Illinois: Institut for Personality and Ability Testing. Questionnaire de personnalité pour enfants, 1968, Montréal: Institut de Recherches psychologiques Inc.

Cattell, R.B. and Warburton, F.W. *Objective Personality and Motivational Tests*. Urbana, Illinois: University of Illinois Press, 1967.

Christiansen, B. *The Aims of the Teaching of Mathematics*. Stockholm: Nordic Committee for the Modernization of School Mathematics, 1967.

Committee for a Comprehensive Education Center. *Proposal for an Experimental Secondary School Research Project to be Conducted Jointly by the New York City Board of Education and the Committee for a Comprehensive Education Center*. New York: Office of High Schools, Board of Education. Unpublished, 1970.

Connes, B. *Course Component Evaluation, Part II: A Dynamic Computer System for Item Banking*. Sherbrooke, Quebec: Centre de Recherches en Psycho-Mathématique, Université de Sherbrooke, 1973.

Cook, W.W., Leeds, C.H., and Callis, R. *Minnesota Teacher Attitude Inventory*. New York: The Psychological Corporation, 1951.

Cook, W.W., Leeds, C.H., and Callis, R. *Inventaire Minnesota des Opinions de l'Enseignant*. Montréal: Institut de Recherches Psychologiques, 1972.

Cronbach, L. Course improvement through evaluation. *Teachers College Record*, *64*, 1963.

Cronbach, L.J. How can instruction be adapted to individual differences. In R.M. Gagné (ed.): *Learning and Individual Differences*. Columbus, Ohio: Charles E. Merrill, 1967.

Cronbach, L.J. and Meehl, P.E. Construct validity in psychological tests. *Psychological Bulletin*, *52*, 1955.

Cronbach, L.J., Rajaratnam, N., and Gleser, G.C. Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology*, *16*, 1963.

Crutchfield, R.S. and Covington, M.L. *A Child Attitude Inventory for Problem Solving (CAIPS)*. Berkeley, Cal.: University of California Department of Psychology, 1972.

Culler, E. and Girden, E. The learning curve in relation to other psychometric functions. American Journal of Psychology, 64, 1951.

Danzinger, K. Fostering creativity: Social and psychological aspects. In D.W. Brison (ed.): Accelerated Learning and Fostering Creativity. Toronto: The Ontario Institute for Studies in Education, 1968.

Davis, R.B. The Madison Project's approach to a theory of instruction. Journal of Research in Science Teaching, 2, 1964.

Day, H.I. Exploratory behaviour as a function of individual differences and level of arousal. Ph.D. Thesis, University of Toronto, 1965. Microfilm #66-1064. Ann Arbor, Michigan: University Microfilms, Inc.

Day, H.I. The effects of increased arousal on attention in high- and low-anxious subjects. Ontario Journal of Educational Research, 9 (3), 1967.

Day, H.I. Anxiety, curiosity, and arousal. Ontario Psychological Association Quarterly, 24, 1967a.

Day, H.I. Evaluations of subjective complexity, pleasingness, and interestingness for a series of random polygons varying in complexity. Perception and Psychophysics, 2 (7), 1967b.

Day, H.I. A subjective definition of complexity. Perceptual and Motor Skills, 25, 1967c.

Day, H.I. The role of specific curiosity in school achievement. Journal of Educational Psychology, 59 (10), 1968.

Day, H.I., Berlyne, D.E., Hunt, D.E. (eds.): Intrinsic Motivation: A New Direction in Education. Toronto: Holt, Rinehart, and Winston of Canada, 1971.

Day, H.I. and Langevin, R. Two necessary conditions for creativity. Journal of Special Education, 3, 1969.

Day, H.I. and Llewellyn, T.E. Effects of amphetamine on selective attention. Perceptive Motivational Skills, 1967.

Decarie, T.G. Intelligence and Affectivity in Early Childhood. New York: International Universities Press, 1965.

DeCharms, R., Morrison, H.W., Reitman, W., and McClelland, D.C. Behavioural correlates of directly and indirectly measured achievement motivation. In D.C. McClelland (ed.): Studies in Motivation. New York: Appleton-Century-Crofts, Inc., 1955.

Demangeon, M. Eprueves d'anxiété. Paris: Institut National d'Etude du Travail et d'Orientation Professionnelle, Conservatoire National des Arts et Metiers, 1966a.

Demangeon, M. Etude de l anxiété en milieu scolaire chez des enfants de 10 - 12 ans. Bulletin de l'Institut National d'Etude du Travail et d'orientation professionnelle, 22(2), 1966b.

Demangeon, M. Une recherche sur l'anxiété par la comparison des groupes appariés. Bulletin de l'Instutut National d'Etude du Travail et d'orientation professionnelle, 27, 1971.

Denny, M.R. Learning. In R. Heber and H. Stevens (eds.): Review of Research in Mental Retardation. Chicago: University of Chicago Press, 1963.

Dienes, Z.P. Concept Formation and Personality. Leicester: Leicester University Press, 1959a.

Dienes, Z.P. Building Up Mathematics. London: Hutchinson Educational, 1960.

Dienes, Z.P. An Experimental Study of Mathematics Learning. London: Hutchinson and Co., 1963.

Dienes, Z.P. Modern Mathematics for Young Children. Harlow (Essex), England: Educational Supply Association, Ltd., 1965.

Dienes, Z.P. (ed.): A Report on Some Current Work in the Problems of Primary Mathematics Learning. Palo Alto, California: The International Study Group for Mathematics Learning, 1965.

Dienes, Z.P. and Jeeves, M.A. Thinking in structures. Psychological Monographs on Cognitive Processus. Vol. I. London: Hutchinson Educational, Ltd., 1965.

Dienes, Z.P. Building Up Mathematics, 3rd ed. London: Hutchinson and Co., 1966.

Dienes, Z.P. and Golding, E.W. I - III: First Years in Mathematics. Harlow (Essex), England: Educational Supply Association, 1966.

Dienes, Z.P. and Golding, E.W. Approach to Modern Mathematics. New York: McGraw-Hill, 1967

Dienes, Z.P. Introducing the Elements of Mathematics. New York: Herder and Herder, 1969.

Dienes, Z.P., Gaulin, C., and Lunkenbein, D. Un programme de mathématique pour le niveau élémentaire (1ère partie). Bulletin de l'association mathématique du Québec, automne - hiver, 1969.

Dienes, Z.P. and Jeeves, M.A. The effects of structural relations on transfer. Psychological Monographs on Cognitive Processes. Vol. II. London: Hutchinson Education, Ltd., 1970.

Dienes, Z.P., Klein S. and Lant, R. Color Secret and Shape Secret. Tests for Measuring Learning Ability. Sherbrooke: Centre de Recherches en Psycho-Mathématique, Université de Sherbrooke, 1973b

Dollard, J. and Miller, N.E. Personality and Psychotherapy. New York: McGraw-Hill, 1950.

Dressel, P.L. The measurement and evaluation of instructional objectives. In D.A. Payne and R.F. McMorris (eds.): Educational and Psychological Measurement. Toronto: Blaisdell, 1967.

Dressel, P.L. and Mayhew, L.B. General Education. Washington, D.C.: American Council on Education, 1954.

Drevdahl, J.E. Factors of importance for creativity. Journal of Clinical Psychology, 12, 1956.

Drevdahl, J.E., & Cattell, R.B. Personality and creativity in artists and writers. Journal of Clinical Psychology, 14, 1958.

Driscoll, J.M. and Lanzetta, J.T. Effects of problem uncertainty and prior arousal on pre-decisional informational search. Psychological Report, 14, 1964.

DuBois, Ph.H. The design of correlations studies in training. In R. Glaser (ed.): Training Research and Education: Science Education. New York: John Wiley and Sons, Inc., 1965.

Duncan, C.P. Recent research on human problem solving. Psychological Bulletin, 56, 1959.

Duncanson, J.P. Intelligence and the ability to learn. Ph.D. Dissertation. Princeton University, 1964.

Dunham, J.L., Guildford, J.P., and Hoepfner, R. Abilities pertaining to classes and the learning concepts. Reports from the Psychological Laboratory, University of Southern California, 39, 1966.

Dunham, J.L., Guilford, J.P., and Hoepfner, R. Multivariate approaches to discovering the intellectual components of concept learning. Psychological Review, 75, 1968.

Dutton, W.H. Measuring attitudes toward arithmetic. Elementary School Journal, 55, 1954.

Dutton, W.H. and Blum, M.P. The measurement of attitudes toward arithmetic with a Likert-type test. The Elementary School Journal, 68, 1968.

Dyer, H.S. The discovery and development of educational goals. Proceedings of the 1966 Invitational Conference on Testing Problems. Princeton, N.J.: Educational Testing Service, 1966.

Ebel, R.L. The relation of testing programs to educational goals. In W.G. Findley (ed.): The Impact and Improvement of School Testing Programs: The 62nd Yearbook of the National Society for the Study of Education. Part II. Chicago: The University of Chicago Press, 1963.

Eggleston, J.F. A Critical Review of Assessment Procedures in Secondary School Science (with special reference to SE). Leicester: University of Leicester Press, 1965.

Eisner, E.W. Educational objectives: Help or hindrance. The School Review, 75, 1967.

Engelhart, M.D. and Beck, J.M. The improvement of tests. In W.G. Findley (ed.): The Impact and Improvement of School Testing Programs: The 62nd Yearbook of the National Scoeity for the Study of Education. Part II. Chicago: The University of Chicago Press, 1963.

Evans, G.T. Attitudes, ends, and means. In B. Crittenden (ed.): Means and Ends in Education. Comments on Living and Learning. Occasional Papers 2. Toronto: The Ontario Institute for Studies in Education, 1969.

Evans, K.M. Sociometry and Education. London: Routledge and Kegan Paint, 1962.

Eysenck, H.H. Reminiscence, drive, and personality theory. Journal of Abnormal Social Psychology, 53, 1965.

Feather, N.T. The relationship of persistence at a task to expectation of success and achievement-related motives. Journal of Abnormal Social Psychology, 63, 1961.

Ferguson, G.A. On learning and human ability. Canadian Journal of Psychology, 8, 1954.

Ferguson, G.A. On transfer and the abilities of men. Canadian Journal of Psychology, 10, 1956.

Findley, W.G. The problem of prediction: Two views. A Paper read at the American Personnel and Guidance Convention. Denver: 1961.

Fink, M. Self-concept as it relates to academic underachievement. California Journal of Education, 13, 1962.

Flanagan, J.C. The uses of educational evaluation in the development of programs, courses, instructional materials, and equipment: Instructional and learning procedures, and administrative arrangements. In R.W. Tyler (ed.): Educational Evaluation: New Roles, New Means: The 68th Yearbook of the National Society for the Study of Education. Part II. Chicago: The University of Chicago Press, 1969.

Flanders, N.A. Teacher influence, pupil attitudes, and achievement. Cooperative Research Monograph, No. 12. Washington, D.C.: U.S. Office of Education, 1965.

Flavell, J.H. The Development of Psychology of Jean Piaget. Princeton, N.J.: Van Nostrand, 1963.

Fleishman, E.A. Human abilities and verbal learning. In R.M. Gagné (ed.): Learning and Individual Differences: A Symposium of the Learning Research and Development Center. University of Pittsburgh. Columbus, Ohio: Charles E. Merrill Publishing Co., 1967.

Flournoy, F. The development of arithmetic understanding tests for primary and intermediate levels. The Journal of Educational Research, 62, 1968.

Fogelman, K.R. Modern maths and intelligence tests. Educational Research, 10, 1968.

Ford, L.H. and Meisels, M. Social desirability and the semantic differential. Educational and Psychological Measurement, 24, 1965.

Foshay, A.W. Educational Achievements of Thirteen-Year-Olds in Twelve Countries. Hamburg, Germany: UNESCO Institute for Education, 1962.

Forrai, T. Iskolai teljesitmény és szorongás (School achievement and anxiety). Budapest: Akadémiai Kiadó, 1968.

Frank, J.D. Recent studies of level of aspiration. Psychological Bulletin, 38, 1941.

Freedman, J.I. Increasing creativity by free-association training. Journal of Experimental Psychology, 1965, 69, 81-91.

Freeman, F.S. Theory and Practice of Psychological Testing. New York: Holt, Rinehard, and Winston, 1962.

Gagné, R.M. The acquisition of knowledge. Psychological Review, 69, 1962.

Gagné, R.M. Learning and proficiency in mathematics. The Mathematics Teacher, 56, 1963.

Gagné, R.M. The analysis of instructional objectives for the design for instruction. ~~In R.M.~~ Glaser (ed.): Teaching Machines and Programmed Learning, II: Data and Directions. Washington, D.C.: National Education Association, 1965.

Gagné, R.M., Mayor, J.R., Garstens, H.L., and Paradise, N.E. Factors in acquiring knowledge of a mathematical task. *Psychological Monographs, No. 526, 76*, 1962.

Gagné, R.M. and Paradise, N.E. Abilities and learning sets in knowledge acquisition. *Psychological Monographs*, *No. 518, 75,* 1961.

Games, P.A. A factorial analysis of verbal learning tasks. *Journal of Experimental Psychology*, *63*, 1962.

Gardner, E.F. and Thompson, G. *Syracuse Scales of Social Relations*. New York: Harcourt, Brace, and World, 1950.

Gaudry, E. and Champion, R.A. The identification of learning-curve parameters with intelligence and motivation. *Australian Journal of Psychology*, *14*, 1962.

Geréb, Gy. *Kisérletek a fáradtság lélektanának köréből*. (Experiments on the psychology of fatigue). Budapest: Adadémaiai Kiadó, 1962,

Geréb, Gy. *Az iskola pszichés klimájáról* (On the psychological climate of the school). Budapest: Tankönyvkiadó, 1970a.

Geréb, Gy. *Intenziv tevékenység és monoton állapot kölcsönhatasanak vizsgálata általános iskolai tanulóknál* (Examination of the interrelation of intensive activity and monotony in elementary students). In *MTA Pszichológiai Tanulmányok 12*. Budapest: Adadémiai Kiadó, 1970b.

Geréb, Gy. and Szabó, Z. *Módszertani elgondolas 10-14 éves tanulók szorongásos álmainak értelmezésére* (A methodological concept for interpreting the anxious dreams of 10-14 year old students). *Magyar Pszichológiai Szemle*, *22*, 1965.

Getzels, J.W and Jackson, P.W. *Creativity and Intelligence: Exploration with gifted students*. New York: Wiley, 1962.

Getzels, J.W. Non-I.Q. intellectual and other factors in college admission. *The Coming Crisis in the Selection of Students for College Entrance*. Washington, D.C.: American Educational Research Association, 1970.

Glaser, R. Evaluation of instruction and changing educational models. In M.C. Wittrock and D.E. Wiley (eds.): *The Evaluation of Instruction: Issues and Problems*. New York: Holt, Rinehart, and Winston, Inc., 1970.

Glennon, V.J. and Callahan, L.G. *Elementory School Mathematics: A Guide to Current Research*, Washington, D.C.: Association for Superivsion and Curriculum Development, NEA, 3rd Edition, 1968.

Gordon, W.J.J. On Being Explicit About the Creative Process. *Journal of Creative Behavior.* *6, (4)*, 1972.

Green, C.W. The relationship between intelligence as determined by intelligence tests and the ability to learn as determined by performance on learning tests. *Journal of Educational Research*, *47*, 1953.

Grimes, J.W. and Allinsmith, W. Compulsivity, anxiety, and school achievement. *Merrill-Palmer Quarterly*, *7*, 1961.

Gronlund, NE. *Sociometry in the Classroom*. New York: Harper Bros., 1959.

Gross, O. *Die Cerebrale Sekundärfunktion*. Leipzig: 1902; cited by Spearman, 1927.

Grossman, B. and J. Wrighter. The relationship between selection-rejection and intelligence, social status, and personality amongst sixth grade children. *Sociometry*, *11*, 1966.

Guilford, J.P. *Psychometric Methods* (2nd ed.). New York: McGraw-Hill, 1954.

Guilford, J.P. Three faces of intellect. *American Psychologist*, *14*, 1959.

Guilford, J.P. Factorial angles to psychology. *Psychological Review*, *68*, 1961.

Guilford, J.P. Intelligence: 1965 model. *American Psychologist*, *21*, 1966.

Guilford, J.P. The nature of human intelligence. New York: McGraw-Hill, 1967a.

Guilford, J.P. Some theoretical views of creativity. In Helson, H. and Beval, W., *Contemporary approaches to psychology*, Van Nostrand Company, 1967b.

Guilford, J.P. *Intelligence, Creativity, and Their Educational Implications*. San Diego: Robert Knapp Publisher, 1968.

Guilford, J.P. Comments on Professor Bloom's paper. In M.C. Wittrock and D.E. Wiley (eds.): *The Evaluation of Instruction: Issues and Problems*. New York: Holt, Rinehart, and Winston, Inc., 1970.

Guilford, J.P. and Hoepfner, R. *The Analysis of Intelligence*. New York: McGraw-Hill, 1971.

Gulliksen, H. *Theory of Mental Tests*. New York: John Wiley and Sons, Inc. (6th printing), 1967.

Guthke, J. Die psychodiagnostische Relevanz von Tests mit Pädagogisierungsphase. *Berichte des 2. Kongresses der Gesellschaft für Psychologie in der DDR*. Berlin: 1969a

Guthke, J. Lernfähigkeit und Leistungsdiagnostik. *Probleme und Ergebnisse der Psychologie*, *27*, 1969b.

Guthke, J. Über die Bedeutung von Lerntests bei der Begabungsdiagnostik. *Probleme und Ergebnisse der Psychologie*, *35*, 1970.

Guthke, J. *Zur Diagnostik der intellektuellen Lernfähigkeit*. Berlin: VEB Deutscher Verlag der Wissenschaften, 1972.

Hadley, S.T. A school mark: fact or fancy? *Educational Administration and Supervision*, *49*, 1954.

Hajtman, B. *Bevezetés a matematikai statisztikába pszichológusok számára. (Introduction to mathematical statistics for psychologists)*. Budapest: Akadémiai Kiadó, 1968.

Hall, C.S. Intercorrelations of measures of human learning. *Psychological Review*, *43*, 1936.

Halliday, A.M. and Redfearn, J.W.I. An analysis of the frequencies of finger tremor in healthy subjects. *Journal of Physiology*, *134*, 1956.

Hamel, M. and Lunkenbein, D. *La Mathématique à l'école élémentaire renouvelée. Vol. I. Les ensembles*. Montréal: Education Nouvelle, 1970.

Hammer, E.F. *Creativity*: *An exploratory investigation of the personalities of Gifted adolescent artists*. New York: Random House, 1961.

Harman, H.H. *Modern Factor Analysis (2nd ed.)*. Chicago: The University of Chicago Press, 1967.

Harris, Ch.W. (ed.): *Problems in Measuring Change*. Madison: The University of Wisconsin Press, 1963.

Haufmann, E. and Kasanin, J. A method for the study of concept formation. *Journal of Psychology*, *3*, 1937.

Hawkins, C.F. and Lanzetta, J.T. Uncertainty, importance, and arousal as determinants of pre-decisional information search. *Psychological Reporter*, *17*, 1965.

Haywood, H.C. Relationships among anxiety, seeking of novel stimuli and level of unassimilated percepts. Journal of Perceptual Learning, 29, 1961.

Haywood, H.C. Novelty-seeking behavior as a function of manifest anxiety and psychological arousal. Journal of Perceptual Learning, 30, 1962.

Haywood, H.C. and Hunt, J. McV. Effects of epinephrine upon novelty preference and arousal. Journal of Abnormal Social Psychology, 67, 1963.

Hebb, D. Drives and the C.N.S. Psychological Review, 62, 1955.

Hebron, M.E. A factorial study of learning a new number system and its relation to attainment, intelligence and temperament. British Journal of Educational Psychology, 32 38 - 45, 1962.

Heckhausen, H. Hoffnung und Furcht in der Leistungsmotivation. Meisenheim/Glan, Hain, 1963.

Heese, K.W. A general factor in improvement with practice. Psychometrika, 7, 1942.

Hemphill, J.K.G. The relationships between research and evaluation studies. In R.W. Tyler (ed.): Educational Evaluation: New Roles, New Means. The 68th Yearbook of the National Society for the Study of Education, Part II. Chicago: The University of Chicago Press, 1969.

Henry, J. Culture against man. New York : Random House, 1963.

Hieronymus, A.N. and Lindquist, E.F. Iowa Tests of Basic Skills. Levels Edition, Form 5. Boston: Houghton Mifflin Co., 1971.

Hively II, W., Patterson, H.L., Page, S.H. A "universe-defined" system of arithmetic achievement tests. Journal of Educational Measurement. 5, (4), 1968.

Hoffman, L.W. and Hoffman, M.L. Review of Child Development Research. Vol. 2. New York: 1966

Hohn, F.E. Teaching creativity in mathematics. Arithmetic Teacher, 52, 1961

Horváth, L.G. Szenzoros és motoros teljesitő-képesseg fáradtság hatására létrejött elváltozásának kisérletes összehasonlitó vizsgálata (An experimental study of modification of sensory and motor efficiency and fatigue). MTA Pszichológiai Tanulmányok 5. Budapest: Akadémiai Kiadó, 1963.

House, B.J. and Zeaman, D. Visual discrimination and intelligence in defectives of low mental age. American Journal of Mental Deficiency, 65, 1960.

Hovland, C.I., Lumsdaine, A.A., and Sheffield, F.D. Experiments on mass communication. Princeton, N.J. Princeton University Press, 1949,

Hughes, J.B., Sprague, J.L., and Bendig, A.W. Anxiety level, response alternation, and performance in serial learning. Journal of Psychology, 38, 1954.

Hull, C.L. Principles of Behavior. New York: Appleton-Century-Crofts, 1943.

Hull, W.P. Concept work with young children. Bulletin of the International Study Group for Mathematics Learning, 1, 1963.

Humphreys, L. Investigations of the Simplex. Psychometrika, 25, 1960.

Hungerman, A.D. Achievement and attitude of sixth-grade pupils in conventional and comtemporary mathematics programs. The Arithmetic Teacher, 14, 1967

Hunt, D.E. Tests of Conceptual Learning. Toronto: Ontario Institute for Studies in Education, 1970a.

Hunt, D.E. Matching Models in Education. The coordination of teaching methods with student characteristics. Toronto: Ontario Institute for Studies in Education. Monograph Series, No.10, 1970b.

Hunt, J. McV. Intelligence and Experience. New York: Ronald Press, 1961

Hunt, J. McV. Toward a history of intrinsic motivation. In H.I. Day, D.E. Berlyne, and D.E. Hunt (eds.): Intrinsic Motivation: A New Direction in Education. The Ontario Institute for Studies in Education, 1971.

Hunt, J. McV., Uzgiris, I. and Wachs, T. Cognitive development in infants of different age levels. Paper Presented at the Meeting of the Society for Research in Child Development. New York: 1967.

Husband, R.W. Intercorrelations among learning abilities. I. Journal of Genetic Psychology, 55, 1939.

Husband, R.W. Intercorrelations among learning abilities. III. Journal of Genetic Psychology, 58, 1941a.

Husband, R.W. Intercorrelations among learning abilities. IV. Effects of age and spread of intelligence upon relationships. Journal of Genetic Psychology, 58, 1941b.

Husén, T. et al (eds) International Study of Achievement in Mathematics: A comparison of Twelve Countries (2 vols., N.Y.; John Wiley and Sons, 1967).

Husén, T. "International Impact of Evaluation" In R.W. Tyler (ed.) Educational Evaluation: New Roles, New Means. The 68th Yearbook of the National Society for the Study of Education. Part II: The University of Chicago Press: Chicago, Illinois, 1969.

Jones, L.V. "Analysis of Variance in its Multivariate Developments." In: Cattell, R.B. (ed.): Handbook of Multivariate Experimental Psychology. Chicago: Rand-McNally and Company, 1966.

Jones, P.S. The Role of Problems in Secondary School Mathematics. In: The Role of Axiomatics and Problem Solving in Mathematics. Boston: Ginn, 1966.

Klein, S., Gondolatok az, alkotóképességröl és vizsgálatának egyes problemairol (Reflections on creativity and some problems of its measurement). Magyar Pszichologiai Szemle, 4, 508-515, 1968.

Klein, S., Titkosírás. Teszt az információelsajátítás és alkotó alkalmazás vizsgálatára. ("Secret Writing" A test of information acquisition and creative application). Unpublished, 1969.

Klein, S., An experiment to develop a new type of intelligence test. Psychology in practice 17, Akadémiai Kiadó, Budapest, 1970,

Klein, S., Kisérlet egy újtipusu intelligencia-teszt kialakítására. ("Experiment to Develop a New Type of Intelligence Test"). Budapest: Akadémiai Kiadó, 1970a.

Klein, S., "Information acquisition and creative application (IA), as a measure of intelligence." Proceedings of the third Congress of the Hungarian Psychological Association. Budapest: Akadémiai Kiadó, 81-82, 1970b.

Klein, S., "As informácioélsajátitás és alkotó alkalmazás (IA) mérése oktatógép segitségével" ("Measuring information acquisition and creative application with a teaching machine"). Magyar Pszichológiai Szemle, 3-4, 1972a.

Kogan, N., and Wallach, M., Risk taking. New York: Holt, Rinehart and Winston, Inc., 1964.

Kónya, A., Influence de l'enseignement spécial sur la persévération. Relation entre la persévération et la créativité. Annales Universitatis Scientiarum Budapestinensis de Rolando Eotuos Nominatae, Paedagogica et Psychologica, 2. 77-85. 1971.

Kramer, K., "Arithmetic Achievement in Iowa and the Netherlands", Elementary School Journal, 59, 258-263, 1959.

Krathwohl, D.R., Stating Objectives Appropriately for Program, for Curriculum, and for Instructional Materials Development , Journal of Teacher Education, 12, 83-92, 1965.

Krathwohl, D.R., Bloom, B.S., and Nasia, B.B., Taxanomy of Educational Objectives. The Classification of Educational Goals. Handbook II. Affective Domain. New York: David McKay Comp. 1964.

Kress, G.C., and Gropper, G.L., The comparison of two strategies for individualizing fixed-pace programmed instruction. American Educational Research Journal, 3, 273-280, 1966.

Kun, M., and Szegedi, M., Az intelligencia mérése (Measurement of Intelligence). Budapest: Akadémiai Kiadó, 1971.

Lamon, W.E. (ed.) Learning and the Nature of Mathematics. Science Research Associates Inc. Chicago, 1972.

Langevin, R. A Study of Curiosity, Intelligence, and Creativity, unpublished PHD. thesis, Toronto: University of Toronto, 1970.

Laurendeau, M. and Pinard, A. Causal thinking in the Child. New York: International Universities Press, 1962.

Lehmann, G. Praktische Arbeits-physiologie. Stuttgart, Georg Thieme Verlag, 1953.

Lerch, H.H. and Hamilton, H. A Comparison of a Structured-Equation Approach to Problem Solving with a Traditional Approach, School Science and Mathematics, 66, 241-246, 1966.

Levy, L.H. Conceptions of Personality. Toronto: Random House, 1970.

Linquist, E.F. The Impact of Machines on Educational Measurement. In: Tyler R.W. (ed.) Educational Evaluation: New Roles, New Means. The Sixty-eight Yearbook of the National Society for the Study of Education Part II. Chicago, Illinois: University of Chicago Press, 1969.

Lindvall, C.M. and Cox R.C. The Role of Evaluation in Programs for Individualized Instruction. In: R.W. Tyler (ed.) Educational Evaluation: New Roles, New Means. The 68th Yearbook of the National Society for the Study of Education Part II. Chicago, Illinois: The University of Chicago Press, 1969.

Lipman, R. S. Learning: Verbal, perceptual-motor, and classical conditioning. In N. R. Ellis (ed.), Handbook of mental deficiency. New York: McGraw-Hill Book Company, 1963.

Lippitt, R. Watson, J. and Westley, B. Planned Change: A comparative study of principles and techniques. New York: Harcourt, Brace and World, Inc. 1958.

Lockard, J. D. (ed). Eighth Report of the International Clearing House on Science and Mathematics Curricular Developments, 1972. Maryland: Science Teaching Center, University of Maryland, 1972.

Loftus, E. J. F. An analysis of the structural variables that determine problem-solving difficulty on a computer-based teletype, Technical Report No. 162. Psychology Series. Stanford, California: Institute for Mathematical studies in the Social Sciences, Stanford University, 1970.

Lord, F. M. and Novick, M. R. Statistical Theories of Mental Test Scores. London: Addison-Wesley Publication Comp., 1968.

Lumsdaine, A. A. Comments on Professor Glaser's Paper In: M. C. Wittrock and D. E. Wiley (eds.). The Evaluation of Instruction. Issues and Problems. New-York: Holt, Rinehart and Winston, Inc, pp..91.95, 1970.

Lyda, W.J. and Church, R.S. Direct, Practical Arithmetical Experiences and Success in Solving Realistic Verbal "Reasoning"Problems in Arithmetic. The Journal of Educational Research, 57, 530-533, 1964.

Lyness, R. C. The Problem as Central in Secondary School Mathematics. Teaching Mathematics: Main Courses in Colleges of Education. Nelson, Lancs.: Coulton, 1967.

Mackinney, J. C. An educational application of a two-dimensional socio-metric test. Sociometry, 356-367, 1948.

MacKinnon, D.W. The highly effective individual. Teachers College Record. 61, 367-378, 1960.

MacKinnon, D.W. The personality correlates of creativity: A study of American architects. In G. Nielson (Ed.) Proceedings of the XIV International Congress of Applied Psychology, Vol. 2. Personality Research. Copenhagen: Munksgaard, 1961.

MacKinnon, D.W. The nature and nurture of creative talent. American Psychologist, 17, 484-495, 1962.

Meyer, W.U., Heckhausen, H. und Kemmler, L. Validierungskorrelate der inhaltsanalytisch erfassten Leistungsmotivation guter and schwacher Schüler des dritten Schuljahres. Psychologische Forschung, 28, 301-326, 1965.

Miel, A.- Berman, L.,ed. In the minds of men: Educating the young people of the world. Report of the world conference on Education at Asilomar, California, March 5-14, 1970. Washington: Association for Supervision and Curriculum Development, NEA, 1970.

Minor, C.A. and Neel, R.G. The relationship between achievement motive and occupational preference. Journal of Counseling Psychology, 5, 39-43, 1958.

Miron, S.M. and Osgood. Ch. E. Language Behavior: the Multivariate Structure of Qualification In: Cattell R.B., (ed.) Handbook of multivariate Experimental Psychology. 1966.

Mitchell, C. How are Pupil's Self-Evaluations? Clearing House, 19, 486-488, 1945.

Muller-Willis, L. Learning Theories of Piaget and Mathematics Instruction, In: Houstoni W.R. (ed). Improving Mathematics Education for Elementary School Teachers: A Conference Report. East Lansing (Mich. U.S.A.): Michigan State University, 38-43, 1967.

Murray, H.A. Techniques for a systematic investigation of fantasy. Journal of Psychology, 115-143, 1936.

Murray, H.A. Explorations in Personality. London: Oxford University Press, 1938.

Nathan, R. and Manco, G. Test de Nathan et Manco. Manuel, Planches de Situation Scolaire Complémentaires au T.A.T. Paris: Centre de Psychologie Appliquée, 1966.

Neugarten, B.L. Social Class and Friendship Among School Children. American Journal of Sociology. 51, 305-313, 1946.

New York College Entrance Examination Board, Report of the Commission on Mathematics, and Appendices to the Report. New York: New York College Entrance Examination Board, 1959.

Northway, M. L. A Primer of Sociometry. Toronto: University of Toronto Press, 1952.

Northway, M.L. and Weld, L. Sociometric Testing: A Guide for Teachers. Toronto: University of Toronto Press, 1957.

Nunnally, J.C. Educational Measurement and Evaluation. New York: McGraw-Hill, 1964.

Ohlsen, M.M. "Interpretation of Test Scores". In: Findley, W.C. (ed.): The Impact and Improvement of School Testing Programs. The sixty-second Yearbook of the National Society for the Study of Education. Part II. Illinois: The University of Chicago Press, Chicago, 1963.

Osgood, C.E. Semantic Differential Technique in the comparative Study of Cultures. American Anthropology, 66, 171-200, 1964.

Osgood, C.E. Exploration in Semantic Space: A Personal Diary Journal of Social Issues Vol. 27 (4), 1971.

Osgood, C.E., Suci, G.J., and Tannenbaum, P.M. The Measurement of Meaning. Urbana, Illinois: University of Illinois Press, 1957.

Otis, A.S. and Lennon, R.T. Epreuve d'habilité Mentale. Montreal: Institut de Recherches Psychologiques, 1967a.

Otis, A.S. and Lennon, R.T. Mental Ability Test. New York: Harcourt, Brace, and World, 1967b.

Palermo, D.S., Castaneda, A. and McCandless, B.R. The relationship of anxiety of children to performance in complex learning tasks. *Child Development*, *27*, 333-337, 1956.

Penney, R. and McCann, B. The Children's Reactive Curiosity Scale, *Psychological Reports*, *15*, 323-334, 1964.

Philips, B. N., Hindsman, E. and Jennings, E. Influence of intelligence on anxiety and perception of self and others. *Child Development*, *31*, 41-46, 1960.

Phillips, B.N., Martin R.P. and Meyers J. Interventions in Relation to Anxiety in School. In Spielberger, Ch.D. (ed): *Anxiety, Current Trends in Theory and Research Vol. II* London: Academic Press, 1972.

Pickerel, E.W. The differential effect of manifest anxiety on test performance. *Journal of Psychology*, *49*, 43-46, 1958.

Pidgeon, D.A. A Comparative Study of Basic Attainments, *Educational Research*, 50-68, 1958.

Pieron, H. Etalonnage d'un test d'attention. Le test de Toulouse et Pieron. *Bulletin de l'Institut National d'Orientation Professionnelle*, 105-112, 1929.

Pillsbury, W.B. *Attention*. New York: The Macmillan Company, 1908.

Pinard, J. W. Test of perseveration: 1. Their relation to character. *British Journal of Psychology*, *23*, 114-128, 1932.

Pólya, G. On teaching Problem Solving. In: *The Role of Axiomatics and Problem Solving in Mathematics*. Boston: Ginn, 1966.

Porter, R.B. and Cattell, R.B. *Children's Personality Questionnaire*. Champaign, Illinois: Institute for Personality and Ability Testing, 1963.

Porter, R.B. and Cattell, R.B. *Questionnaire de Personnalité pour Enfants*. Paris: Centre de Psychologie Appliquée, 1965.

Provus, M. Evaluation of Ongoing Programs in the Public School System In: R.W. Tyler (ed.) *Educational Evaluation: New Roles, New Means*. *The 68th Yearbook of the National Society for the Study of Education*. Part II. Chicago, Illinois: The University of Chicago Press, 1969.

Rajaratnam, Nagaswari, Cronbach, L.J., & Gleser, Goldine C. Generalizability of stratified parallel tests. *Psychometrika*, *30*, 39-56, 1965.

Raphelson, A.C. The relationship between imaginative, direct, verbal, and physiological measures of anxiety in an achievement situation. *Journal of Abnormal Social Psychology*, *54* 13-18, 1957.

Rapier, J. The learning abilities of normal and retarded children as a function of social class. *Journal of Educational Psychology*, *59*, 101-10, 1968.

Raven, J.C. *Standard Progressive Matrices*. London: H.K. Lewis, 1958.

Reithman, W.R. and Atkinson, J.W. Some methodological problems in the use of thematic apperceptive measures of human motives. In: Atkinson, J.W. (ed.) *Motives in Fantasy, Action, and Society*. Princeton: D. Van Nostrand, 664-684, 1958.

Reymond-Rivier, B. *Choix Sociometriques et Motivations*. Neuchatel: Delachauix et Niestle, 1961.

Reys, R.E. and Post, T.R. *The Mathematics Laboratory*. Theory to Practice. Boston: Prindle, Weber and Schmidt, Inc., 1973.

Ricciuti, H.N. and Clark, R.A. *A comparison of need-achievement stories written by experimentally "relaxed" and "achievement-oriented" subjects: effects obtained with new pictures and revised scoring categories*. Princeton, N.J: Educational Testing Service, 1954.

Riedesel, C.A. Verbal problem solving: Suggestions for improving instruction. *The Arithmetic Teacher*, 1964, *5*, 312-316, 1964.

Rogers, C.R. Toward a theory of creativity. In S.J. Parnes and H.F. Harding (eds) *A source book for creative thinking* New York: Charles Scribuer's Sons 64-72, 1962.

Róna, B., Hegedüs, G., Kapasvári, J., Mága, R., Merétei, K, and Szabó, P. Az egész napos iskola 4. Osztályos tanulóinak komplex vizsgálata. (Complex examination of grade 4 day schoool pupils). In: Szabadkai, S. (ed.) *Tanulmányck a hazai egész napos iskola 1-4. osztályainak gvakorlatából (Studies of the Practices of Hungarian Day Schools in The First 4 Grades)*. Budapest: Tankönyvkiadó Vállalat, 87-103, 1971.

Rosenbloom, P.C. Problem Making and Problem Solving. *The Role of Axiomatics and Problem Solving in Mathematics*. Boston: Ginn, 1966.

Rosenbloom, P.C. (ed.) *Modern Viewpoints in the Curriculum*. New York: McGraw-Hill, 1964.

Roughead, W.G. and Scandura, J.M. What is Learned in Mathematical Discovery. *Journal of Educational Psychology*, *59*, 283-289, 1968.

Ruch, F.L. The method of common points of mastery as a technique in human learning experimentation. *Psychological Review*, *43*, 229-234, 1936.

Ruch, F.L. Measuring gain from a common point of mastery. *Psychol. Rep.*, *9*, 1961.

Ruebush, B.K. Anxiety, In: H.W. Stevenson, J. Kagan and C. Spieker (eds) Child Psychology 62nd N.S.S.E. Yearbook. 460-516, 1963.

Rusch, R.R., Denny, D.A. and Ives, S. Fostering creativity in sixth grade. Elementary School Journal, 65. 262-268, 1965.

Russell, H.- Horn, D.M. (eds.) Experiences with sets and Numbers. Mathematics Evaluation Materials Package Project. Toronto, Ontario: The Ontario Institute for Studies in Education Curriculum Series 14., 1972.

Sanford, R.N. Will psychologists study human problems? American Psychologist, 20, 192-202, 1965.

Sarason, I.G. Empirical findings and theoretical problems in the use of anxiety scales. Psychological Bulletin, 57, 403-415, 1960.

Sarason, S.E., Davidson, K. Lighthall, F. and Waite, R. Classroom observation of high and low anxious children. Child Development, 29, 277-286, 1958.

Sarason S. Davidson K., Lightfall F. Waite R. Ruebush B. Anxiety in Elementary School Children. John Wiley and Sons, 1960,

Sarason, S.B., Mandler, G. and Craighill, P.G. The Effect of Differential Instructions on Anxiety and Learning Journal of abnormal and social Psychology, 47, 561-565, 1952.

Scandura, J.M. Problem solving and prior learning. Journal of Experimental Education, 34, 7-11, 1966a.

Scandura, J.M. Prior learning, presentation order, and presentation order, and and prerequisite practice in problem solving. Journal of Experimental Education, 34, 12-18, 1966b.

Scandura, J.M. An analysis of exposition and discovery modes of problem solving instruction. Journal of Experimental Education, 33, 149-160, 1964.

Scandura, J.M. (ed.) Research in Mathematics Education. Washington: A Special Publicaition of the National Council of Teachers of Mathematics. 1967.

Schröder, H. Kombinierter Lern- und Intelligenztest . KLI4+. Deutsche Schultests. Herausgegeben von K. Ingenkamp. Weinheim und Berlin: Julius Beltz, 1968.

Scriven, M. Methodology of Education. American Educational Research Assoc. Monograph Series on Airriculum Evaluation I. Chicago: Rand McNally, 39-83, 1967a.

Scriven, M. "The Methodology of Evaluation. Aera Monograph Series on Curriculum Evaluation, No.1. Chicago: Rand McNally, 39-89, 1967b.

Servais, W. and Varga, T. Teaching School Mathematics. Paris: Unesco, 1971.

Shaw, M.C. Need achievement scales as predictors' of academic success. Journal of Educational Psychology, 52, 282-285, 1961.

Shenathirajah, N. - Weiss, N. Evaluation in Geography. A Resource Book for Teachers. Curriculum Series 10. Toronto, Ontario: The Ontario Institute for Studies in Education, 1971.

Sherman, H. Common Elements in New Mathematics Programs: Origins and Evolution. New York: Teachers College Press, 1972.

Shoobs, N.E. Sociometry in the Classroom, Sociometry, 154-164, 1947.

Simon, B. (ed.) Psychology in the Soviet Union. London: Routledge and Kegan Paul, 1957.

Simpson, R.M. "Creative Imagination" *American Journal of Psychology* *33*, 234-43, 1922.

Sitgreaves, R. Statistical Problems in Curriculum Experimentation and Evaluation. In: Rosenbloom, P.C. (ed.) *Modern Viewpoints in the Curriculum Experimentation*, September 25-28, 1961. New York: McGraw-Hill, 157-163, 1964.

Skinner, B.F. *Science and Human Behavior*. New York: Macmillan Co., 1953.

Skinner, B.F. *Verbal Behavior*. New York: Appleton- Century - Crofts, 1957.

Smock, C.D. The influence of psychological stress on the "intolerance of ambiguity." *J. abnorm. Soc. Psychol*. *50*, 1955a.

Smock, C.D. The influence of stress on the perception on incongruity. *J. abnormal soc. Psychol*. *50*, 1955b.

Snider J.G, and Osgood Ch. E. *Semantic Differential Technique, a Sourcebook*. Chicago: Aldine Publishing Company, 1969.

Sorenson, A.G. The use of teaching machines in developing an alternative to the concept of intelligence. *Educational and Psychological Measurement*, *23*, 323-329. 1963.

Spearman, C. *The Abilities of Man*. New York: The MacMillan Company, 1927.

Spence, K.W. *Behaviour Theory and Conditioning*. New Haven: Yale University Press, 1956.

Spielberger, C.D. The effects of anxiety on complex learning and academic achievement. In: Spielberger, C.D. (ed.) *Anxiety and Behaviour*. New York: Academic Press, 1966.

Spielberger, C.D. (ed.) Anxiety and Behaviour. New York: Academic Press, 1966 (a) 2nd ed. 1968.

Spielberger, C.D. (ed.) Anxiety. Current trends in theory and research Vol. I and II. New York and London: Academic Press, 1972.

Stake, R.E. Learning parameters, aptitudes, and achievements. Psychometric Monogram, No. 9, 1961.

Stake, R.E. The countenance of educational evaluation . Teachers College Record. 68 (7) 1967.

Stake, R.E. and Denny, T. Needed Concepts and Techniques for Utilizing More Fully the Potential of Evaluation. In: R.W. Tyler (ed.) Educational Evaluation, New Roles, New Means. The 68th Yearbook of the National Society for the Study of Education. Part II Chicago, Illinois: University of Chicago Press, 1969.

Stanley, J.C. Reliability . In: Thorndike, R.L. (ed.) Educational Measurement. Washington, D.C.: American Council on Education, 1971.

Starch, D. and E.C. Elliott, Relability of Grading Work in Mathematics. School Review, 21, 254-259, 1913.

Start, K.B. Teacher Assessment as a problem of sampling within a complex composite criterion. International Review of Applied Psychology. 21 (1), 53-60, 1972.

Steffe, L.P. The relationship of conservation of numerousness to problem-solving abilities of first-grade children. The Arithmetic teacher, 15, 47-52, 1968.

Stier, L.D. A comparison of self-evaluation, peer evaluation, teacher evaluation and standardized tests on personality and achievement factors in the fifth and eighth grades. Unpublished Ph. D. dissertation. Berkeley: University of California, 1948.

Stufflebeam, D.L. A depth study of the evaluation. Theory into Practice. 121-133, 1966.

Sullivan, E.V. *Piaget and the School Curriculum. A Critical Appraisal*. Bulletin No.2 Toronto: The Ontario Institute for Studies in Education, 1967.

Surányi, J., Reményi, K, and Varga, T. *Principles of updating school mathematics as seen in Hungary*. Background paper of Working Conference on Mathematics Learning, Budapest, 1972. Manuscript.

Taylor, C.W. (ed.) *Creativity: Progress and Potential*. New York: McGraw-Hill, 1964.

Taylor, J.A. The relationship of anxiety to the conditioned eyelid response. *Journal of Experimental Psychology*, *41*, 81-92, 1951.

Taylor, J.A. A personality scale of manifest anxiety. *Journal of Abnormal Social Psychology*, *48*, 285-290, 1953.

Taylor, J.A. The effects of anxiety level and psychological stress on verbal learning. *Journal of Abnormal Social Psychology*, 57, 55-60, 1958.

Taylor, J.B. "What do attitude scales measure: The problem of social desirability. *Journal of Abnormal and Social Psychology*, *62 (2)*, 386-390, 1961.

Taylor, P.A. and Maguire, T.O. "A theoretical evaluation model". *The Manitoba Journal of Educational Research*, *1*, 12-17, 1966.

Thompson, G.G. and Witryol, S.L. The relationship between intelligence and motor learning ability. Journal of Psychology, *22*, 237-246, 1946.

Thorndike, R.L. (ed.) *Educational Measurement*. Second Ed. Washington D.C.: American Council on Education, 1971.

Thurstone, L.L., and Chave, E.J. *The Measurement at Attitude*, Chicago: University of Chicago Press, 1929.

Thurstone, L.L. The learning function. Journal of Genetic Psychology, 3, 469-497, 1930.

Tilton, J.W. Intelligence test scores as indicative of ability to learn. Educational and Psychological Measurement, 9, 291-296, 1949.

Tichener, E.B. Lectures on the Elementary Psychology of Feeling and Attention. New York: The Macmillan Company, 1908.

Torrance, E.P. (ed.). Rewarding Creative Thinking. Minneapolis, Minnesota: University of Minnesota, 1960.

Torrance, E.P. Measurement and development of the creative thinking Abilities. Gifted Child. Year Book of Education Edited by G. Z.F. Bereday and J.-A. Lauwerys. (Ed.) London: Evan Bros., Ltd. 1962a.

Torrance, E.P. Guiding Creative Talent. Englewood Cliffs, New Jersey: Prentice-Hall, 1962b.

Torrance, E.P. Education and creativity. In Taylor, C.P. (ed.) Creativity: Progress and Potential. New York, McGraw-Hill, 1964.

Torrance, E.P. Torrance Tests of Creative Thinking. Norms-Technical Manuals, Test booklets. Princeton, N.J.: Personnel Press, Inc., 1966.

Torrance, E.P. Torrance Tests of Creative Thinking. Direction Manuals and Scoring Guide. Form A Verbal Tests, Form B. Verbal and Figural Tests. Princeton, N.J.: Personnel Press, Inc., 1968.

Torrance, E.P. *Teaching skills for creative ways of learning*. Chicago: Science Research Associates 1969.

Torrance, E.P. *Torrance Tests of Creative Thinking*. Directions Manuel and Scoring Guide. Form A Figural test, Princeton, N.J.: Personnel Press, Inc., 1970.

Torrance, E.P. *Tests de Pensée Créative et Manuel*, Paris: Centre de Psychologie Appliqué, 1972.

Traub, R.E. Some implications for educational evaluation. In: Crittenden, B. (ed.): *Means and Ends in Education*. *Comments on Living and Learning*. Occasional Papers 2. Toronto: The Ontario Institute for Studies in Education, 1969.

Traxler, A.E. and North, R.D. The selection and use of tests in a school testing program. In: Findley, W.G. (ed.): *The Impact and Improvement of school testing Programs*. *The Sixty-Second Yearbook of the National Society for the Study of Education* Part II. Chicago, Illinois: The University of Chicago Press, 1963.

Tucker, L. Determination of generalized learning curves by factor analysis. In: Dubois, P.H., Manning, W.L. and Spies, C.J. (eds.): *Factor analysis and related techniques in the study of learning*. ONR contract 816602, Technical Report, No. 7. St. Louis: Washington University, 1959.

Tuddenham, R.D. A 'Piagetian' test of cognitive development. In: Dockrell, W.B. (ed.): *On Intelligence. Contemporary Theories and Educational Implications: A Symposium. Toronto, 1969*. Toronto: The Ontario Institute for Studies in Education, 1970.

Tyler, R.W. *Constructing Achievement Tests*. Columbus, Ohio: Ohio State University, 1934.

Tyler, R.W. General statement on evaluation. *Journal of Educational Research*, 492-501, 1942.

Tyler, R.W. (ed.) *Educational Evaluation: New Roles, New Means. The 68th yearbook of the National Society for the Study of Education.* Part II. The University of Chicago Press: Chicago, Illinois 1969.

Uhlinger, C.A. and Stephens, M.W. Relation of achievement motivation to academic achievement in students of superior ability. *Journal of Educational Psychology*, *51*, 259-266, 1960.

Varga, T. The use of a composite method for the mathematical education of young children. *Bulletin of the International Study Group for Mathematics Learning*, *3* (2) 1-9, 1965.

Varga, T. Higher mathematics in lower grades. In: Feldman, L. (ed.) *Prepubescent Mathematical Learning*. San José, California: San José State College Mathematics Department, C18-C22, 1968.

Varga, T. *Beszámoló a komplex matematikatanitási kisérletről.* (Report about the Composite Method of Mathematics Teaching). Budapest: Országos Pedagógiai Intézet, 1970. Manuscript.

Vernon, M.D., *Human Motivation*. Cambridge: University Press, 1969.

Vygotsky, L.S. *Thought and Language*; New York: Wiley, 1962.

Walbesser, H.H. *An Evaluation Model and Its Application: Second Report*, Washington, D.C.: The American Association for the Advancement of Science. AAAS Miscellaneous Publication 68-4, 1968.

Wallach, M.A. and Kogan, N. *Modes of Thinking in Young Children. A Study of the Creativity-Intelligence Distinction*. New York: Holt, Rinehart and Winston, 1965.

Warburton, F.W. The British Intelligence Scale. In: Dockvell, W.B. (ed.): *On intelligence*. *Contemporary theories and educational implications: A symposium: Toronto, 1969.* Toronto: The Ontario Institute for Studies in Education, 1970.

Weaver, J.F. Evaluation and the Classroom Teacher. In: Begle, E.G. (ed.) *Mathematics Education*. *The Sixty-ninth yearbook of the National Society for the Study of Education Part I*. The University of Chicago Press, 1970.

Wechsler, D. *Mannal, Wechsler Intelligence Scale for Children*, New York: The Psychological Corporation, 1949.

Weinstein, G. - Fantini, M.D. *Toward Humanistic Education*. *A Curriculum of Affect*. New York: Praeger Publishers, 1970.

Weisberg, P.S. and Springer K.J. Environmental factors in creative function: A study of gifted children. *Archives of general Psychiatry*. 1961, *5*, 554-564.

Weiss, P., Wertheimer, M. and Groesbeck, B. Achievement motivation, academic aptitude, and college grades. *Educational Psychological Measurement*, 1959, *19*, 663-666.

Wendt, H.W. Motivation, Effort, **and Performance**. In: McClelland, D.C. (ed.) *Studies in Motivation* New York: Appleton-Century-Crofts, Inc. 448-459, 1955.

Wesley, E.L. Perseverative behaviour in a concept formation task as a function of manifest anxiety and rigidity. *Journal of Social Psychology*, 1953, *48*, 129-134.

Whitla, D.K. "Research in College Admissions" *Educational Evaluation: New Roles, New Means*. *68th yearbook of the National Society for the Study of Education*, Part II. Chicago: University of Chicago Press, 1969.

Wiersma, E. Die Sekundarfunction bei Psychosen. Journal fur Psychologie und Neurologie, 1906, 8, 1-24.

Wiley, L.E. and Wiley, A.M. Studies in learning function Psychometrika, 1937, 2, 107-120.

Williams, J.D. Research on the Dienes method by the National Foundation for Educational Rseearch in England and Wales., Bulletin of the International Study Group for Mathematics Learning, 1963, 1 (1) 3-9.

Williams, J.D. Some problems of design. Bulletin of the International Study Group for Mathematics Learning, 1965, 3, (4) 19-37.

Williams, J.D. Evaluation in education: Ten dilemmas of design. The Statistician, 1967, 17 (2) 183-191.

Williams, J.D. Problems and possibilities in the assessment and investigation of mathematics learning. Educational Studies in Mathematics, 1971, 4, 135-149.

Williams, J.D. Concept Formation and the Design of the Learning Environment. (A "research system" for discussion at the Florence I.S.G.M.L. Meeting). Centre de Recherches en Psycho-Mathematique, Sherbrooke, Quebec, 1972a, Unpublished.

Williams, J.D. "Modern mathematics, intelligence and basic conceptual skills", Journal of Structural Learning, 3(2), 1972b.

Williams, J.D. Course-component Evaluation Part I: Rationale and Outline. Sherbrooke, Quebec: Centre de Recherche en Psycho-Mathematique, Université de Sherbrooke, 1973. Manuscript.

Williams, J.D. and Fogelman, K.R. "Sets, logic and intelligence tests." Journal of Structural Learning, 1972, 3, (3) 25-36.

Willoughby, S.S. "Who Won the International Contest?" The Arithmetic Teacher, 1968, 15, 623-629.

Wilson, F.S. Stuckey, T. Langevin, R. Are Pupils in the Open Plan School Different? The Journal of Educational Research, 66 (3) 115-118, 1972.

Wilson, J.W. The Role of Structure in Verbal Problem Solving. The Arithmetic Teacher, 1967, 14, 486-497.

Witkin, H.A., Lewis, H.B.,Hertzman, M., Machover, K., Meissner, P.B., Wapner, S., Personality Through Perception, New York: Harper and Row, 1954.

Witkin, H.A.,Dyak, R.B., Faterson, H.F., Goodenough, D.R., and Karp, S.S. Psychological Differentiation, New York: John Wiley and Sons, Inc., 1962.

Wittrock, M.C. The evaluation of instruction: Cause-and-Effect Relations in Naturalistic Data. In: Wittrock, M.C. and Wiley, D.E. (eds.) The Evaluation of Instruction. Issues and Problems. New York: Holt, Rinehart and Winston, Inc., 3-21, 1970.

Wittrock, M.C. and Wiley, D.E. (eds.) The Evaluation of Instruction. Issues and Problems. New York: Holt, Rinehart, and Winston, Inc., 1970.

Wood, R. Objectives in the Teaching of Mathematics. Educational Research, 1968, 10, 83-98.

Woodrow, H. The relation between abilities and improvement with practice. Journal of Educational Psychology, 1938, 29, 215-230.

Woodrow, H. The application of factor-analysis to problems of practice. Journal of General Psychology, 21, 457-460, 1939a.

Woodrow, H. The common factors in fifty-two mental tests. Psychometrika, 4, 99-108, 1939b.

Woodrow, H. Factors in Improvement with practice. Journal of Psychology, 7, 55-70, 1939c.

Worthen, B.R. A Study of Discovery and Expository Presentation: Implications for Teaching. The Journal of Teacher Education, 19, 223-242, 1968.

Wundt, W. Outlines of psychology. Leipzig: Engelmann, 1897.

Wyatt, R.H. A study of power spectra analysis of normal finger tremors. IEEE Transactions on Bio-Medical Engineering, 15, (1) 33-45, 1968.

Yamamoto, K. Experimental Scoring Manuals for Minnesota Tests of Creative Thinking and Writing. Kent, Ohio: Bureau of Educational Research Kent State University 1964.

Yeager, J.L. and Lindvall, C.M. An Exploratory Investigation of Selected Measures of Rate of Learning. The Journal of Experimental Education, 36, 78-81, 1967.

Zazzo, R. and Stambak, M. Test de persévération. In: Zazzo, R. Manuel pour l'Examen Psychologique de l'Enfant. II. 3rd Ed. Neuchâtel: Delachaux et Niestlé, 1969.

Zeaman, D. & House, Betty J. The relation of I.Q. and learning. In Gagné R.M. (Ed.), Learning and individual differences. Columbus, Ohio: Merrill, 192-217, 1967.